# The Numerical Solution of Diffusion Problems in Strongly Heterogeneous Non-Isotropic Materials

James Hyman, Mikhail Shashkov
Los Alamos National Laboratory
T-7, MS-B284
Los Alamos NM 87545
and
Stanly Steinberg
Department of Mathematics and Statistics
University of New Mexico
Albuquerque NM 87131

# Contents

1	Intr	Introduction							
2	The	Properties of the Continuum Problem	$\epsilon$						
3	The Spaces of Discrete Functions								
	3.1	The Discretization of Scalar and Vector Functions	8						
		3.1.1 The Discrete Scalar Functions	11						
		3.1.2 The Discrete Vector Functions	11						
	3.2	The Spaces of Discrete Functions	12						
		3.2.1 The Space of Discrete Scalar Functions	12						
		3.2.2 The Space of Discrete Nodal Vector Functions	13						
		3.2.3 The Space of Discrete Surface Vector Functions	14						
		3.2.4 The Formal and Natural Inner Products	16						
4	The	Finite-Difference Method	19						
	4.1	The Prime Operator	19						
		4.1.1 The Nodal Discretization for Vectors	20						
		4.1.2 The Surface Discretization for Vectors	20						
	4.2	The Derived Operator	21						
		4.2.1 The Surface Discretization for Vectors	21						
		4.2.2 The Nodal Discretization for Vectors	22						
	4.3	The Discrete Operator Equations	23						
	4.4	The Discrete Operators on a Rectangular Grid	24						
		4.4.1 The Cell-Node Discretization	24						
		4.4.2 The Cell-Surface Discretization	25						
	4.5	Theoretical Properties of the Algorithms	26						
		4.5.1 The Null Space of the Operator $\mathcal{G}$	27						
	4.6	Solving the System of Linear Equations	27						
		4.6.1 Solving for the Fluxes	28						
5	Nur	nerical Examples	29						
	5.1	Piecewise Linear Solutions	31						
	5.2	Piecewise Quadratic Solutions	32						
	5.3	Non-Diagonal Continuous K	34						
	5.4	Non-Diagonal Discontinuous <b>K</b>	35						
	5.5	Diagonal Discontinuous K: Homogenization Example	37						
	5.6	Flow Through a Sand-Shale System	40						
	5.7	Flow Through a System Containing an Impermeable Streak	42						
	5.8	Non-Diagonal, Piecewise Continuous K	43						
A	The	Truncation Error for the Discrete Analog of K grad	47						

# **Abstract**

A new second-order finite-difference algorithm for the numerical solution of diffusion problems in strongly heterogeneous and non-isotropic media is constructed. On problems with rough coefficients or highly nonuniform grids, the new algorithm is superior to all other algorithms we have compared it with. For problems with smooth coefficients on smooth grids, the method is comparable with other second order methods. The new algorithm is formulated for logically-rectangular grids and is derived using the *support-operators method*.

A key idea in deriving the method was to replace the usual inner product of vector functions by an inner product weighted by the inverse of the material properties tensor and to use the flux operator, defined as the material properties tensor times the gradient, rather than the gradient, as one of the basic first-order operators in the support-operators method. The discrete analog of the flux operator must also be the negative adjoint of the discrete divergence, in an inner product that is a discrete analog of the continuum inner product. The resulting method is conservative and the discrete analog of variable coefficient Laplacian is symmetric and negative definite on nonuniform grids. In addition, on any grid, the discrete divergence is zero on constant vectors, the null space for the gradient are the constant functions and, when the material properties are piecewise constant, the discrete flux operator is exact for piece-wise linear functions. We compare the methods on some of the most difficult examples to be found in the literature.

# 1 Introduction

The main goal of this paper is the description and investigation of a new finite-difference algorithms for solving the elliptic partial differential equation (PDE) or stationary diffusion equation

$$-\mathbf{div}\,\mathbf{K}\,\mathbf{grad}\,u = f\,,\quad (x,y) \in V\,. \tag{1.1}$$

The solution, u = u(x, y), is the concentration to be solved for (temperature in heat diffusion problems, and pressure in flow problems). Here V is a two dimensional region,  $\operatorname{div}$  is the divergence,  $\operatorname{grad}$  is the gradient,  $\mathbf{K} = \mathbf{K}(x, y)$  is a symmetric positive-definite matrix, and f = f(x, y) is a given right-hand side or forcing function. The boundary conditions are general Robin (or mixed):

$$(\mathbf{Kgrad}\ u,\ \vec{n}) + \alpha u = \psi\ , \quad (x,y) \in \partial V\ , \tag{1.2}$$

where  $\vec{n}$  is vector of unit outward normal to the boundary  $\partial V$ , and  $\alpha$  and  $\psi$  are functions given on  $\partial V$ . The algorithm is constructed using a nontrivial generalization of the *support-operators method* for solving problems where the material properties tensor (or matrix)  $\mathbf{K}$  may be discontinuous and non-diagonal and, moreover, the computational grid may not be smooth.

The support-operators method constructs discrete analogs of invariant differential operators div and grad, which satisfy discrete analogs of the integral identities responsible for the conservative properties of the continuum model. The method was initially developed in [14] by Samarskii, Tishkin, Favorskii, and Shashkov and is fully described in [17].

This paper is the third of a series on the support operators method. In the first paper [16], the support operators method was combined with the mapping method to produce an algorithm for equations with general boundary conditions. The resulting method was shown to be accurate when both  $\mathbf{K}$  is smooth and the problem is solved on a smooth grid. In the second paper [15], the support-operators method was extended to define new cell-centered finite-difference algorithm for solving time-dependent diffusion equations with discontinuous diagonal  $\mathbf{K}$  on logically rectangular non-smooth grids, such as the grids associated with Lagrangian hydrodynamics calculations. This paper also contains an extensive review of the literature on constructing approximations of differential operators on non-uniform grids and motivation for using the support operators method.

In this paper we extend the support-operators method to non-diagonal non-smooth tensor K and non-smooth logically-rectangular grids. A key to improving the accuracy for non-smooth K is to use the flux operator K grad, rather than the gradient operator grad, as one of the basic first-order operators. This requires the usual inner product of vector functions be replaced by an inner product weighted by the inverse of the material properties tensor. The method is linear, conservative and material discontinuities are assumed to occur at the surfaces of the grid cells. The method uses both the heat flux and the temperature as primary variables.

The discrete analog of the variable-coefficient Laplacian div K grad can be decomposed as a composition of two discrete operators: a divergence DIV; and a flux operator

-KGRAD, which are the adjoints of each other. This ensures the self-adjointness and negative definiteness of the discrete variable-coefficient Laplacian for general grids. Moreover, on any grid, the discrete analog of div is exactly equal to zero only on constant vectors, the grad is equal to zero only for constant functions and (when the material properties are piece-wise constant) the discrete flux operator is exact for piecewise linear functions. On rectangular grids, all discrete the operators reduce to standard finite difference approximations and when the discontinuous heat conductivity is a scalar, the algorithm produces the appropriate harmonic average heat conductivity for the fluxes.

Because the fluxes are the primary unknowns for the method, when this approach is used to solve (1.1), the new method has twice as many unknowns as the more traditional algorithms. However, because the method is second-order accurate on general logically rectangular grids, fewer grid points are needed than the commonly used first order methods to give the same level of accuracy in the solution. In our computational comparisons with many of the existing methods, when the computational mesh is significantly skewed, we find that the new method is much more accurate than the traditional finite difference and finite volume methods. Also, because the discrete difference equations are symmetric and positive definite, only half the coefficients for the difference equations need be stored and the equations can be solved with some of the most powerful iterative methods for solving linear systems.

This paper is arranged as follows. In Section 2, the continuum problem is written in operator form to illuminate the properties of the operators that should have analogs in the discrete case. The main point in this section is to introduce an inner-product on vectors that is weighted by  $\mathbf{K}^{-1}$  and to introduce the operator  $\mathbf{K}$  grad rather than just grad. In Section 3, the grid and the discretizations of scalar and vector functions are given. Both nodal and surface discretizations for fluxes are introduced, and the discrete inner product for general non-diagonal  $\mathbf{K}$  is constructed.

In Section 4, following the support-operators method, approximations for div and K grad are derived using both the nodal and surface discretizations. Using these operators, the finite difference scheme for div K grad is constructed. The theoretical properties of the discrete operators are summarized (see Appendix A also), and it is shown that, for the surface discretization, the null-space of grad are the constant functions, while for the nodal discretization, the null-space also contains the spurious highest-frequency mode on a square grid.

In Section 4.6, we describe strategies for solving the linear equations given by the discretizations of the variable-coefficient Laplacian. An important point here is that the discrete operator for the surface discretization is not local. However, it is the product of a local discrete operator with the an inverse of a local operator, so the residual can still be computed as a local operation. Iterative methods that only require local operations, such as the preconditioned conjugate gradient method, can be used efficiently.

In Section 5, we test and compare our algorithms on eight of the most difficult examples to be found in the literature. These examples verify that the surface discretization approach performs reliably on all of the examples, and the nodal discretization gives reasonable results.

The first example confirms that the surface discretization method is exact when the solution is piecewise linear. The second example demonstrates the second-order convergence rate for discontinuous diagonal K. The third example demonstrates the second-order convergence rate for both the surface and nodal discretizations for continuous non-diagonal K. The fourth example demonstrates the second-order convergence for the surface discretization and first-order convergence for the nodal discretization for non-diagonal and discontinuous K.

The next three examples are related to porous-media flows. The fifth example compares the discretizations on a highly-discontinuous (checkerboard)  $\mathbf{K}$ . The surface discretization is significantly more accurate and less noisy than the nodal discretization. The sixth example simulates the random placement of blocks of shale in sand, and shows that the surface discretization produces far more accurate results than the nodal discretization. The seventh example involves an curved low-permeability streak in a porous media. A non-uniform grid is adapted to the streak and  $\mathbf{K}$  is non-diagonal. The simulation confirms that the flow field conforms to the low-permeability region.

The eighth example combines all of the difficult features of the previous problems. We verify that for a non-diagonal and piecewise constant K, the surface discretization is exact on a non-uniform grid. We also verify that the nodal discretization is second-order accurate.

In Appendix A, we prove the surface discretization K grad is exact for piecewise linear functions when K is piecewise constant.

# 2 The Properties of the Continuum Problem

In this section, we will develop the flux form of the elliptic PDE (1.1) as a system of first-order equations and analyze the system in terms of abstract operators on inner-product spaces of scalar and vector-valued functions. The analysis will be for Robin (mixed) boundary conditions (1.2). The case with Dirichlet boundary conditions is a straight forward extension of this analysis and will not be analyzed here.

We introduce the space of scalar functions H with the inner product

$$(u,\,v)_H = \int_V u\,v\,dV + \oint_{\partial V} u\,v\,dS\,,\quad u,v\in H\,,$$
 (2.1)

and rewrite Equations (1.1,1.2) as

$$\mathbf{A} u = \mathbf{F} . \tag{2.2}$$

The operator A is given by

$$\mathbf{A}: H \to H, \quad \mathbf{A}u = \begin{cases} -\mathbf{div} \, \mathbf{K} \, \mathbf{grad} \, u, & (x,y) \in V \\ (\mathbf{K} \, \mathbf{grad} \, u, \, \vec{n}) + \alpha \, u, & (x,y) \in \partial V \end{cases} \tag{2.3}$$

and has the following properties:

$$(\mathbf{A}u, v)_H = (u, \mathbf{A}v)_H, \quad (\mathbf{A}u, u)_H > 0,$$
 (2.4)

The right-hand side of (2.2) has the form

$$\mathbf{F} = \begin{cases} f, & (x,y) \in V \\ \psi, & (x,y) \in \partial V \end{cases}$$
 (2.5)

We investigate the properties of problem (1.1,1.2) by writing it in terms of first-order operators in flux or mixed form:

$$\mathbf{div}\,\vec{w} = f, \quad (x,y) \in V,$$

$$\vec{w} = -\mathbf{K}\,\mathbf{grad}\,u, \quad (x,y) \in V,$$

$$-(\vec{w},\vec{n}) + \alpha\,u = \psi, \quad (x,y) \in \partial V.$$

$$(2.6)$$

The flux  $\vec{w} = \mathbf{G} u = -\mathbf{K} \operatorname{\mathbf{grad}} u$  has physical meaning and is continuous across discontinuities in K, but grad u has neither. When the matrix K is non-diagonal and discontinuous, there are advantages (and it is natural) to analyzing the equations by considering the operator G = -K grad directly, rather than the operator grad and matrix K separately.

To investigate (2.6), we introduce the space of vector functions H with the inner product of two vector functions  $\vec{A}$ ,  $\vec{B} \in \mathbf{H}$ , defined by

$$(\vec{A}, \vec{B})_{\mathbf{H}} = \int_{V} (\mathbf{K}^{-1} \vec{A}, \vec{B}) dV.$$
 (2.7)

Because the matrix K is symmetric and positive definite, so is  $K^{-1}$  and (2.7) satisfies all the axioms of an inner product. This "weighted" inner product is well defined for discontinuous  ${f K}$  and naturally arises in mixed finite-element formulations (see, for example, [5, 22]).

From (2.6), it is clear that operator A can be represented in the form

$$\mathbf{A} = \mathbf{\Omega} + \mathbf{D} \cdot \mathbf{G} \,, \tag{2.8}$$

where the operators G, D, and  $\Omega$  have the following definitions

$$\mathbf{G}u = -\mathbf{K} \operatorname{\mathbf{grad}} u, \quad (x, y) \in V, \tag{2.9}$$

$$\mathbf{D}\vec{w} = \begin{cases} +\mathbf{div}\,\vec{w}\,, & (x,y) \in V\,, \\ -(\vec{w},\vec{n})\,, & (x,y) \in \partial V\,, \end{cases}$$
(2.10)

$$\mathbf{D}\vec{w} = \begin{cases} -\mathbf{K} \mathbf{g} \mathbf{I} \mathbf{d} u, & (x,y) \in V, \\ -\mathbf{d} \mathbf{i} \mathbf{v} \vec{w}, & (x,y) \in V, \\ -(\vec{w}, \vec{n}), & (x,y) \in \partial V, \end{cases}$$

$$\mathbf{\Omega}u = \begin{cases} 0, & (x,y) \in V, \\ \alpha u, & (x,y) \in \partial V. \end{cases}$$

$$(2.10)$$

Here

$$G: H \to H ; D: H \to H ; \Omega: H \to H .$$
 (2.12)

Using the first-order operators, System (2.6) can be rewritten in the form

$$\mathbf{\Omega} u + \mathbf{D} \vec{w} = F , \quad \vec{w} = \mathbf{G} u . \tag{2.13}$$

A crucial relation which we must retain in our discrete approximation is

$$\mathbf{D} = \mathbf{G}^* \,. \tag{2.14}$$

This is clear from the definition of operator  $\mathbf{D}$ , the definition (2.1) for the inner product in the space H, and integral identity

$$\int_{V} \phi \operatorname{\mathbf{div}} \vec{w} \, dV + \int_{V} (\vec{w}, \operatorname{\mathbf{grad}} \phi) \, dV = \oint_{S} \phi \left( \vec{w}, \vec{n} \right) dS \,, \tag{2.15}$$

which give:

$$(\mathbf{D}\,\vec{w},\,u)_{H} = \int_{V} u\,\mathbf{div}\,\vec{w}\,dV - \oint_{\partial V} u\,(\vec{w},\,\vec{n})\,dS$$

$$= -\int_{V} (\vec{w},\,\mathbf{grad}u)\,dV$$

$$= -\int_{V} (\vec{w},\,\mathbf{K}^{-1}\,(\mathbf{K}\,\mathbf{grad}u)\,dV \qquad (2.16)$$

$$= (\vec{w},\,\mathbf{G}\,u)_{\mathbf{H}}.$$

Also, it is evident that  $\Omega = \Omega^* \geq 0$ .

Because  $\mathbf{A} = \mathbf{\Omega} + \mathbf{G}^* \mathbf{D}$ , the properties (2.4) follow from the properties of operators  $\mathbf{\Omega}$ ,  $\mathbf{D}$  and  $\mathbf{G}$ . Note that boundary conditions are included in definitions of operators and spaces of functions in natural way. The properties of first-order operators discussed in this Section are preserved by the finite-difference methods derived using the support-operators method.

# 3 The Spaces of Discrete Functions

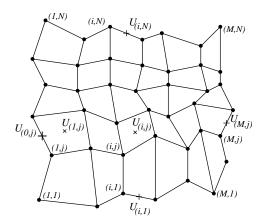
In this section, we define our notation for a logically rectangular grid [9], a cell-centered discretization of scalar functions, and both nodal and face-centered discretizations of vector-valued functions.

## 3.1 The Discretization of Scalar and Vector Functions

The nodes of a logically-rectangular grid can be indexed the same way as a rectangular grid with indices (i,j),  $1 \le i \le M$ ,  $1 \le j \le N$  (see Figure (3.1)). The quadrangle defined by the nodes (i,j), (i+1,j), (i+1,j+1), and (i,j+1) is called the (i,j) cell (see Figure 3.2). The area of this cell is denoted by  $VC_{(i,j)}$ . The length of the side of the (i,j) cell that connects the vertices (i,j) and (i,j+1) is denoted  $S\xi_{(i,j)}$ , while the length of the side that connects the vertices (i,j) and (i+1,j) is denoted  $S\eta_{(i,j)}$ . The angle between any two adjacent sides of cell (i,j) that meet at node (k,l) is denoted  $\varphi_{k,l}^{(i,j)}$  (the angle  $\varphi_{(i+1,j)}^{(i,j)}$  is displayed in Figure 3.2). We assume, unless otherwise stated, that the cells are convex. (Meshes with non-convex cells are considered in Section 3.2.2 and [15].)

To study convergence rates, we impose some standard mild smoothness assumptions on the family of grids. A small parameter which characterize the density of the grid is

$$h = \max\left\{\frac{1}{M-1}, \frac{1}{N-1}\right\}. \tag{3.1}$$



 $Figure \ 3.1: \ A \ Logically \ Rectangular \ Grid \ and \ the \ Discretization \ of \ a \ Scalar$ 

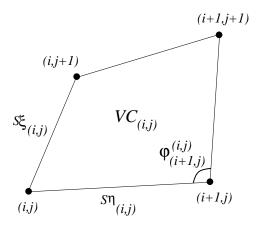


Figure 3.2: A Typical Cell of a Logically Rectangular Grid

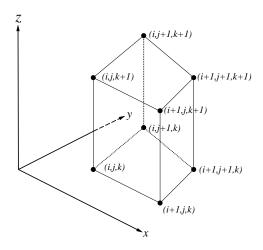


Figure 3.3: The 3-D Mesh

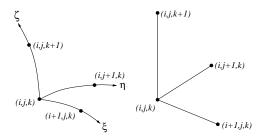


Figure 3.4: Curvilinear Coordinates and Grid Lines

We assume that there exits constants  $C_{max}^{(1)}$  and  $C_{min}^{(1)}$ , which do not depend on h, such that

$$C_{min}^1 h^2 \le V C_{(i,j)} \le C_{max}^1 h^2 ,$$
 (3.2)

and that exists constants  $C_{max}^{(2)}$  and  $C_{min}^{(2)}$ , which do not depend on h, such that

$$C_{min}^2 h \le S\xi_{(i,j)}, S\eta_{(i,j)} \le C_{max}^2 h,$$
 (3.3)

and that there exits a constant  $\epsilon > 0$ , which does not depend on h, such that

$$\sin\left(arphi_{(k,l)}^{(i,j)}
ight) \ge \epsilon \,. \tag{3.4}$$

Our notation is motivated by considering the 2-D grid as a projection of a 3-D grid. This approach may seem awkward at first, but it becomes natural when put into a three dimensional setting and clarifies how the finite-difference methods generalize to 3-D. Because in 2-D the functions depend only on the two coordinates x and y, we first introduce a grid

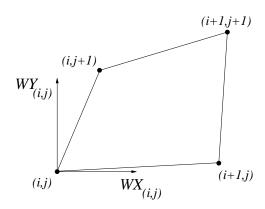


Figure 3.5: The Nodal Discretization of a Vector

in z in the third dimension. The z grid is chosen to construct a right prism of height one for each mesh point of grid in the plane for which the cell in the plane is the base (see Figure 3.3). It is also useful to interpret the grid as the discretization of a map from a curvilinear coordinate system  $x = x(\xi, \eta, \zeta)$ ,  $y = y(\xi, \eta, \zeta)$ ,  $z = z(\xi, \eta, \zeta)$ , where the nodes of the grid are given by  $x_{(i,j,k)} = x(\xi_i, \eta_j, \zeta_k)$ , with  $\xi_i = i \Delta \xi$ , and so forth, as shown in Figure 3.4. Thus increasing i corresponds to increasing  $\xi$  and so on. Using this 3-D interpretation, the notation  $S\xi_{(i,j)}$  refers the area of the 3-D surface given by the points (i,j,k), (i,j+1,k), (i,j,k+1), (i,j+1,k+1), that is,  $S\xi$ , gives the element of surface area for a surface where  $\xi$  is constant, because we took the height of the prism equal to one. Similar results hold for other sides of a 2-D cell.

#### 3.1.1 The Discrete Scalar Functions

The discrete analog of the scalar function u is the cell-centered discrete scalar function  $U_{(i,j)}$  (see Figure 3.1), whose indices vary in the same range as the volume  $VC_{(i,j)}$ . The treatment of the boundary conditions requires the introduction of the values of the scalar function on the centers of the boundary segments (see Figure 3.1):

$$U_{(0,j)}, \quad U_{(M,j)}, \quad j=1,\ldots,N-1; \quad U_{(i,0)}, \quad U_{(i,N)}, \quad i=1,\ldots,M-1.$$
 (3.5)

In the 3-D interpretation, scalar functions are defined in the centers of the 3-D prisms and in the centers of the boundary surfaces. Again, because we only consider the 2-D case, these values can be projected to the 2-D cells, and the centers of the boundary sides.

The components of **K** are discretized in the same way as u. The scalar functions  $\alpha$  and  $\psi$  from the boundary conditions, are discretized in the same way as u is on the boundary.

#### 3.1.2 The Discrete Vector Functions

Two possibilities are used for discretizing vector functions  $\vec{W} = (WX, WY)$ : the first uses the usual Cartesian components  $WX_{(i,j)}$  and  $WY_{(i,j)}$  of the vector at the nodes as shown

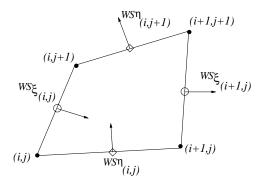


Figure 3.6: The Surface Discretization of a Vector

in Figure 3.5; and the second uses the orthogonal projections of the vector on the direction which is perpendicular to the surfaces of 3-D cells at the centers of the cells. Because the 3-D cell is a right prism, we can interpret these components as the orthogonal projections on the directions which are perpendicular to the sides of the cell, as in Figure 3.6. The notation

$$WS\xi_{(i,j)}: i=1,\ldots,M; j=1,\ldots,N-1$$
 (3.6)

is used for the component at the center of side  $S\xi_{(i,j)}$ , and the notation

$$WS\eta_{(i,j)}: i=1,\ldots,M-1; j=1,\ldots,N$$
 (3.7)

is used for the component at the center of side  $S\eta_{(i,j)}$ . Note that the two components of a vector are defined at the same point.

# 3.2 The Spaces of Discrete Functions

The spaces of discrete scalar and vector functions need inner products. For scalar functions this is straight forward, but for vector functions, there are two spaces:  $\mathcal{HN}$  for the nodal discretization; and  $\mathcal{HS}$  for the surface discretization. Neither inner-product is simple because of the use of  $\mathbf{K}$  in the inner product and, in both cases, the values of  $\mathbf{K}$  are not given at the same points as the vector components.

#### 3.2.1 The Space of Discrete Scalar Functions

The space of discrete scalar functions is labeled HC and has the inner product

$$(U, V)_{HC} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} U_{(i,j)} V_{(i,j)} V C_{(i,j)}$$

$$+ \sum_{i=1}^{M-1} U_{(i,0)} V_{(i,0)} S \eta_{(i,1)} + \sum_{j=1}^{N-1} U_{(N,j)} V_{(N,j)} S \xi_{(N,j)}$$

$$(3.8)$$

$$+ \sum_{i=1}^{M-1} U_{(i,M)} \, V_{(i,M)} \, S \eta_{(i,M)} + \sum_{j=1}^{N-1} U_{(0,j)} \, V_{(0,j)} \, S \xi_{(1,j)} \, .$$

#### 3.2.2 The Space of Discrete Nodal Vector Functions

The space of discrete nodal vector functions is called  $\mathcal{HN}$  and the inner product on this space is given by  $(\cdot, \cdot)_{\mathcal{HN}}$ . Note that  $\mathbf{K}$  and consequently  $\mathbf{K}^{-1}$  are defined at cell centers, but the components the vectors are defined at the cell nodes, which complicates the definition of an inner product. At first, we assume that the components of the vectors  $\vec{A}$  and  $\vec{B}$  are defined at the cell centers, and then a natural inner product in the space vector functions is

$$(\vec{A}, \vec{B}) = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (\mathbf{K}^{-1} \vec{A}, \vec{B})_{(i,j)} VC_{(i,j)}, \qquad (3.9)$$

where  $(\mathbf{K}^{-1} \vec{A}, \vec{B})$  is the inner product of vectors  $\mathbf{K}^{-1} \vec{A}$  and  $\vec{B}$ , and the inner product of two vectors  $\vec{A}$  and  $\vec{B}$  is

$$(\vec{A}, \vec{B}) = AXBX + AYBY. \tag{3.10}$$

For this type of discretization it is natural to assume that tensor K is given by its Cartesian components:

$$K_{xx}, \quad K_{xy} = K_{yx}, \quad K_{yy},$$
 (3.11)

and

$$\mathbf{K}\,\vec{A} = \left(\begin{array}{c} K_{xx}\,AX + K_{xy}\,AY \\ K_{xy}\,AX + K_{yy}\,AY \end{array}\right). \tag{3.12}$$

and consequently

$$(\mathbf{K}^{-1}\vec{A},\vec{B}) = (\mathbf{K}^{-1})_{xx} AX BX + (\mathbf{K}^{-1})_{xy} (AX BY + AY BX) + (\mathbf{K}^{-1})_{yy} AY BY. \quad (3.13)$$

For simplicity, introduce the notation  $KI = K^{-1}$  for the matrix inverse of K.

Because the vectors are defined at nodes, and not at the cell centers, to define the inner product for the space of nodal vector functions  $\mathcal{HN}$ , we must interpolate the vectors. We define

$$(\vec{A}, \vec{B})_{\mathcal{HN}} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (\mathbf{KI}\vec{A}, \vec{B})_{(i,j)} VC_{(i,j)}$$
(3.14)

where the inner product in a cell is given by

$$(\mathbf{K}\mathbf{I}\vec{A}, \vec{B})_{(i,j)} = \sum_{k,l=0}^{1} V_{(i+k,j+l)}^{(i,j)}$$

$$\{(\mathbf{K}\mathbf{I}_{xx})_{(i,j)} \ AX_{(i+k,j+l)} \ BX_{(i+k,j+l)} +$$

$$(\mathbf{K}\mathbf{I}_{xy})_{(i,j)} \left[ AX_{(i+k,j+l)} \ BY_{(i+k,j+l)} + AY_{(i+k,j+l)} \ BX_{(i+k,j+l)} \right] +$$

$$(\mathbf{K}\mathbf{I}_{yy})_{(i,j)} \ AY_{(i+k,j+l)} \ BY_{(i+k,j+l)} \}$$
(3.15)

and the  $V_{(i+k,j+l)}^{(i,j)}$  are weights satisfying

$$\sum_{k,l=0}^{1} V_{(i+k,j+l)}^{(i,j)} = 1.$$
(3.16)

In this formula, each index (k,l) corresponds to one of the vertices of the (i,j) cell, and the notation for the weights is the same as for the angles of the cell. As was shown in [15], to obtain a first-order approximation for gradient operator, it is necessary that the weights  $V_{(i+k,j+l)}^{(i,j)}$  are one half of the area of the triangle in (i,j) cell, which contains angle at the node (i+k,j+l), divided by the volume of the cell  $VC_{(i,j)}$ .

Because the weights are positive if the cell is convex, we are guaranteed that the discrete analog of variable-coefficient Laplacian is positive-definite (see [15] for details). When the cell is not convex, we modify the definition,

$$\tilde{V}_{(i+k,j+l)}^{(i,j)} = \frac{\left|V_{(i+k,j+l)}^{(i,j)}\right|}{\sum_{p,q=0}^{1} \left|V_{(i+p,j+q)}^{(i,j)}\right|},$$
(3.17)

to give positive weights (see [15]). For convex cells, the two definitions coincide.

#### 3.2.3 The Space of Discrete Surface Vector Functions

The space of discrete surface vector functions is called  $\mathcal{HS}$  and the inner product on this space is  $(\cdot, \cdot)_{\mathcal{HS}}$ . Again, there is the problem that the components of the vectors are not defined at the cell centers where  $\mathbf{K}$  is defined. We use the same approach and notation as before for the basic weighted inner product (3.9).

When the vectors are defined at cell centers and  $\mathbf{K} = \mathbf{I}$ , then the formula for the inner product (3.9) in terms of the components of the vectors perpendicular to the cell sides must be defined (see Figure 3.7). Suppose that the  $\xi$  and  $\eta$  axes form a non-orthogonal basis system and that  $\varphi$  is the angle between these axes. If the unit normals to the axes are  $n\vec{S}\xi$  and  $n\vec{S}\eta$ , then the components of the vector  $\vec{W}$  in this basis are the orthogonal projections  $WS\xi$  and  $WS\eta$  of  $\vec{W}$  onto the normal vectors. (See the discussion in Chapter 2 of Knupp and Steinberg [9] for more details.) A simple vector algebra calculation shows that if  $\vec{A} = (AS\xi, AS\eta)$  and  $\vec{B} = (BS\xi, BS\eta)$ , then the expression for the inner product is

$$(\vec{A}, \vec{B}) = \frac{AS\xi BS\xi + AS\eta BS\eta + (AS\xi BS\eta + AS\eta BS\xi) \cos(\varphi)}{\sin^2(\varphi)}.$$
 (3.18)

We now consider the case of a non-diagonal matrix K with the Cartesian components

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{xx} & \mathbf{K}_{xy} \\ \mathbf{K}_{xy} & \mathbf{K}_{yy} \end{pmatrix} . \tag{3.19}$$

Denote by  $\varphi_1$  and  $\varphi_2$ , the angles between the x-axis of Cartesian coordinate system and the first and second axes of local coordinate system, respectively (see Figure 3.8). In terms of

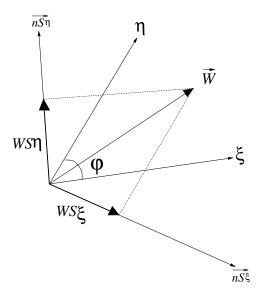


Figure 3.7: The Components of a Vector in a Local Basis

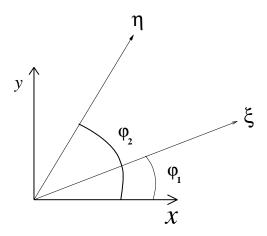


Figure 3.8: The Angles  $\varphi_1$  and  $\varphi_2$ 

the coordinates of vectors in the local basis system, the inner product  $(\mathbf{K}\mathbf{I}\vec{A},\vec{B})$  is

$$(\mathbf{K}\mathbf{I}^{-1}\,\vec{A},\vec{B}) = \frac{T11\,AS\xi\,BS\xi + T22\,AS\eta\,BS\eta + T12\,(AS\xi\,BS\eta + AS\eta\,BS\xi)}{\sin^2(\varphi)}\,,\qquad(3.20)$$

where

$$T11 = \mathbf{K}\mathbf{I}_{xx} \cos^{2} \varphi_{1} + 2 \mathbf{K}\mathbf{I}_{xy} \cos \varphi_{1} \sin \varphi_{1} + \mathbf{K}\mathbf{I}_{yy} \sin^{2} \varphi_{1},$$

$$T12 = \mathbf{K}\mathbf{I}_{xx} \cos \varphi_{1} \cos \varphi_{2}$$

$$+ 2 \mathbf{K}\mathbf{I}_{xy} (\cos \varphi_{1} \sin \varphi_{2} + \sin \varphi_{1} \cos \varphi_{2}) + \mathbf{K}\mathbf{I}_{yy} \sin \varphi_{1} \sin \varphi_{2},$$

$$T22 = \mathbf{K}\mathbf{I}_{xx} \cos^{2} \varphi_{2} + 2 \mathbf{K}\mathbf{I}_{xy} \cos \varphi_{2} \sin \varphi_{2} + \mathbf{K}\mathbf{I}_{yy} \sin^{2} \varphi_{2}.$$

This formula is used to obtain the discrete inner product in the cell:

$$(\mathbf{K}\mathbf{I}\vec{A}, \vec{B})_{(i,j)} = \sum_{k,l=0}^{1} \frac{V_{(i+k,j+l)}^{(i,j)}}{\sin^{2}(\varphi_{(i+k,j+l)}^{(i,j)})}$$

$$\left[T11_{(i+k,j+l)}^{(i,j)} AS\xi_{(i+k,j)} BS\xi_{(i+k,j)} + T22_{(i+k,j+l)}^{(i,j)} AS\eta_{(i,j+l)} BS\eta_{(i,j+l)} + (-1)^{k+l} T12_{(i+k,j+l)}^{(i,j)} \left(AS\xi_{(i+k,j)} BS\eta_{(i,j+l)} + AS\eta_{(i,j+l)} BS\xi_{(i+k,j)}\right)\right],$$
(3.21)

where, for example,

$$T11_{(i+k,j+l)}^{(i,j)} = (\mathbf{KI}_{xx})_{(i,j)} \cos^{2} \left( (\varphi_{1})_{(i+k,j+l)}^{(i,j)} \right) \\ + 2 (\mathbf{KI}_{xy})_{(i,j)} \cos \left( (\varphi_{1})_{(i+k,j+l)}^{(i,j)} \right) \sin \left( (\varphi_{1})_{(i+k,j+l)}^{(i,j)} \right) \\ + (\mathbf{KI}_{yy})_{(i,j)} \sin^{2} \left( (\varphi_{1})_{(i+k,j+l)}^{(i,j)} \right).$$

That is, the values of matrix elements are defined in the cell (i, j), and the angle  $\varphi_1$  is related to the corresponding vertex of the cell. The formulas for T12 and T22 are similar. In general, the notation is the same as for any quantity related to cell and vertex. Finally, as in (3.14), the inner product in  $\mathcal{HS}$  is given by

$$(\vec{A}, \vec{B})_{\mathcal{HS}} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (\mathbf{KI}\vec{A}, \vec{B})_{(i,j)} VC_{(i,j)}.$$
 (3.22)

#### 3.2.4 The Formal and Natural Inner Products

To compute the adjoint relationships, it is helpful to introduce a formal inner products,  $[\cdot, \cdot]$ , in the spaces of scalar and vector functions. In HC

$$[U, V]_{HC} = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} U_{(i,j)} V_{(i,j)}$$

$$+ \sum_{i=1}^{M-1} U_{(i,0)} V_{(i,0)} + \sum_{j=1}^{N-1} U_{(M,j)} V_{(M,j)} + \sum_{i=1}^{M-1} U_{(i,N)} V_{(i,N)} + \sum_{j=1}^{N-1} U_{(0,j)} V_{(0,j)} ,$$
(3.23)

in  $\mathcal{H}\mathcal{N}$ 

$$[\vec{A}, \vec{B}]_{\mathcal{HN}} = \sum_{i=1}^{M} \sum_{j=1}^{N} AX_{(i,j)} BX_{(i,j)} + \sum_{i=1}^{M} \sum_{j=1}^{N} AY_{(i,j)} BY_{(i,j)},$$
(3.24)

and in  $\mathcal{HS}$ 

$$[\vec{A}, \vec{B}]_{\mathcal{HS}} = \sum_{i=1}^{M} \sum_{j=1}^{N-1} AS\xi_{(i,j)} BS\xi_{(i,j)} + \sum_{i=1}^{M-1} \sum_{j=1}^{N} AS\eta_{(i,j)} BS\eta_{(i,j)}.$$
(3.25)

Then the relationships between the natural inner products and the formal inner products are:

$$(U, V)_{HC} = [\mathcal{M} U, V]_{HC}, \quad (\vec{A}, \vec{B})_{\mathcal{H}\mathcal{N}} = [\mathcal{N} \vec{A}, \vec{B}]_{\mathcal{H}\mathcal{S}}, \quad (\vec{A}, \vec{B})_{\mathcal{H}\mathcal{S}} = [\mathcal{S} \vec{A}, \vec{B}]_{\mathcal{H}\mathcal{S}},$$

$$(3.26)$$

where  $\mathcal{M}$ ,  $\mathcal{N}$ , and  $\mathcal{S}$  are matrices.

Formulas for these matrices can be found by direct comparison of the formal and natural inner products. The formula for  $\mathcal{M}$  is

$$(\mathcal{M} \ U)_{(i,j)} = V C_{(i,j)} U_{(i,j)}, \quad i = 1, \dots, M-1; \quad j = 1, \dots, N-1.$$

$$(\mathcal{M} \ U)_{(i,j)} = S \xi_{(i,j)} U_{(i,j)}, \quad i = 0 \text{ and } i = M; \quad j = 1, \dots, N-1.$$

$$(\mathcal{M} \ U)_{(i,j)} = S \eta_{(i,j)} U_{(i,j)}, \quad i = 1, \dots, M-1; \quad j = 0 \text{ and } j = N.$$

$$(3.27)$$

From these formulas, we see that the matrix  $\mathcal{M}$  is a symmetric positive-definite operator in the formal inner product:

$$[\mathcal{M} U, V]_{HC} = [U, \mathcal{M} V]_{HC}, \quad [\mathcal{M} U, U]_{HC} > 0.$$
 (3.28)

The operator  $\mathcal N$  can be written in block form:

$$\mathcal{N}\vec{A} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} \begin{pmatrix} AX \\ AY \end{pmatrix} = \begin{pmatrix} N_{11}AX + N_{12}AY \\ N_{21}AX + N_{22}AY \end{pmatrix}. \tag{3.29}$$

A comparison of the formal inner product  $(\vec{A}, \vec{B})_{HN}$  given in (3.14) and the natural inner product gives

$$[\mathcal{N}\vec{A}, \vec{B}]_{\mathcal{H}\mathcal{N}} = \sum_{i=1}^{M} \sum_{j=1}^{N} \{ [(N_{11} AX)_{(i,j)} + (N_{12} AY)_{(i,j)}] BX_{(i,j)} + (3.30)$$

$$[(N_{21} AX)_{(i,j)} + (N_{22} AY)_{(i,j)}] BY_{(i,j)} \}.$$

Note that all components of operator  $\mathcal{N}$  are diagonal operators,

$$(N_{11} AX)_{(i,j)} = n11_{(i,j)} AX_{(i,j)}, \qquad (N_{12} AY)_{(i,j)} = n12_{(i,j)} AY_{(i,j)}, (N_{21} AX)_{(i,j)} = n21_{(i,j)} AX_{(i,j)}, \qquad (N_{22} AY)_{(i,j)} = n22_{(i,j)} AX_{(i,j)},$$
(3.31)

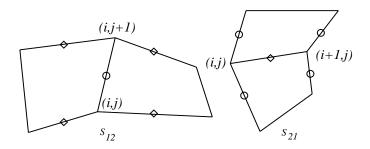


Figure 3.9: The Stencils for the Operators  $S_{12}$  and  $S_{21}$ 

and the diagonal elements are given by

$$n11_{(i,j)} = (\mathbf{K}\mathbf{I}_{xx})_{(i,j)} V_{(i,j)}^{(i,j)} + (\mathbf{K}\mathbf{I}_{xx})_{(i-1,j)} V_{(i,j)}^{(i-1,j)} + (\mathbf{K}\mathbf{I}_{xx})_{(i-1,j-1)} V_{(i,j)}^{(i-1,j-1)} + (\mathbf{K}\mathbf{I}_{xx})_{(i,j-1)} V_{(i,j)}^{(i,j-1)} ,$$

$$n12_{(i,j)} = (\mathbf{K}\mathbf{I}_{xy})_{(i,j)} V_{(i,j)}^{(i,j)} + (\mathbf{K}\mathbf{I}_{xy})_{(i-1,j)} V_{(i,j)}^{(i-1,j)} + (\mathbf{K}\mathbf{I}_{xy})_{(i-1,j-1)} V_{(i,j)}^{(i-1,j-1)} + (\mathbf{K}\mathbf{I}_{xy})_{(i,j-1)} V_{(i,j)}^{(i,j-1)} ,$$

$$n21_{(i,j)} = (\mathbf{K}\mathbf{I}_{xy})_{(i,j)} V_{(i,j)}^{(i,j)} + (\mathbf{K}\mathbf{I}_{xy})_{(i-1,j)} V_{(i,j)}^{(i-1,j)} + (\mathbf{K}\mathbf{I}_{xy})_{(i-1,j-1)} V_{(i,j)}^{(i-1,j-1)} + (\mathbf{K}\mathbf{I}_{xy})_{(i,j-1)} V_{(i,j)}^{(i,j-1)} ,$$

$$n22_{(i,j)} = (\mathbf{K}\mathbf{I}_{yy})_{(i,j)} V_{(i,j)}^{(i,j)} + (\mathbf{K}\mathbf{I}_{yy})_{(i-1,j)} V_{(i,j)}^{(i-1,j-1)} V_{(i,j)}^{(i,j-1)} .$$

$$+ (\mathbf{K}\mathbf{I}_{yy})_{(i-1,j-1)} V_{(i,j)}^{(i-1,j-1)} + (\mathbf{K}\mathbf{I}_{yy})_{(i,j-1)} V_{(i,j)}^{(i,j-1)} .$$

This operator is symmetric and positive-definite in the formal inner product:

$$[\mathcal{N} \vec{A}, \vec{B}]_{\mathcal{H}\mathcal{N}} = [\vec{A}, \mathcal{N} \vec{B}]_{\mathcal{H}\mathcal{N}}, \quad [\mathcal{N} \vec{A}, \vec{A}]_{\mathcal{H}\mathcal{N}} > 0.$$
 (3.33)

The operator S can be written in block form:

$$S\vec{A} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} AS\xi \\ AS\eta \end{pmatrix} = \begin{pmatrix} S_{11} AS\xi + S_{12} AS\eta \\ S_{21} AS\xi + S_{22} AS\eta \end{pmatrix}.$$
(3.34)

The operators  $S_{11}$  and  $S_{22}$  are diagonal and the stencils for the operators  $S_{12}$  and  $S_{21}$  are shown on Figure 3.9. A comparison of natural inner product  $(\vec{A}, \vec{B})_{HS}$  and the formal inner product

$$[S\vec{A}, \vec{B}]_{\mathcal{HS}} = \sum_{i=1}^{M} \sum_{j=1}^{N-1} [(S_{11} AS\xi)_{(i,j)} + (S_{12} AS\eta)_{(i,j)}] BS\xi_{(i,j)}$$

$$+ \sum_{i=1}^{M-1} \sum_{j=1}^{N} [(S_{21} AS\xi)_{(i,j)} + (S_{22} AS\eta)_{(i,j)}] BS\eta_{(i,j)},$$

$$(3.35)$$

gives

$$(S_{11} A S \xi)_{(i,j)} = \left( \sum_{k,l=0}^{1} \frac{V_{(i,j+l)}^{(i-k,j)} T 1 1_{(i,j+l)}^{(i-k,j)}}{\sin^{2}(\varphi_{(i,j+l)}^{(i-k,j)})} \right) A S \xi_{(i,j)},$$

$$(S_{12} A S \eta)_{(i,j)} = \sum_{k,l=0}^{1} (-1)^{k+l} \frac{V_{(i,j+l)}^{(i-k,j)} T 1 2_{(i,j+l)}^{(i-k,j)}}{\sin^{2}(\varphi_{(i,j+l)}^{(i-k,j)})} A S \eta_{(i-k,j+l)},$$

$$(S_{21} A S \xi)_{(i,j)} = \sum_{k,l=0}^{1} (-1)^{k+l} \frac{V_{(i+l,j)}^{(i,j-k)} T 1 2_{(i+l,j)}^{(i,j-k)}}{\sin^{2}(\varphi_{(i+l,j)}^{(i,j-k)})} A S \xi_{(i+l,j-k)},$$

$$(S_{22} A S \eta)_{(i,j)} = \left( \sum_{k,l=0}^{1} \frac{V_{(i+l,j)}^{(i,j-k)} T 2 2_{(i+l,j)}^{(i,j-k)}}{\sin^{2}(\varphi_{(i+l,j)}^{(i,j-k)})} \right) A S \eta_{(i,j)}.$$

We remark that these formulas are valid only for  $i=2,\cdots,M-2;\ j=2,\cdots,N-2$ , but it is easy to show that if fictitious nodes are introduced for  $i=0,\ i=M+1,\ j=0,$  and j=N+1, whose coordinates are the same as for the corresponding real nodes, then the formulas are valid for all i and j. If all the weights  $V_{(i,j)}^{p,q}$  are positive, then the operator  $\mathcal S$  is symmetric and positive-definite in the formal inner product:

$$[\mathcal{S}\vec{A}, \vec{B}]_{\mathcal{HS}} = [\vec{A}, \mathcal{S}\vec{B}]_{\mathcal{HS}}, \quad [\mathcal{S}\vec{A}, \vec{A}]_{\mathcal{HS}} > 0. \tag{3.37}$$

# 4 The Finite-Difference Method

We now use the support operators method to derive a discrete the divergence, flux operator, and variable coefficient Laplacian. We first derive a discrete approximation to the divergence, and then use this discrete divergence to derive the approximations to the flux operator and Laplacian using discrete analogs of the integral identities. Because the principle role of the divergence operator, we call it the *prime operator*. Because the discrete approximations for the flux operator and Laplacian are derived from the prime operator, they are called *derived operators*.

# 4.1 The Prime Operator

A natural conservative invariant definition of the divergence operator is

$$\operatorname{\mathbf{div}} \vec{w} = \lim_{V \to 0} \frac{1}{V} \oint_{V} (\vec{w}, \vec{n}) \, dV \,. \tag{4.38}$$

This identity is used in [17] to derive a discrete analog **DIV** of the divergence **div**, for both the nodal and surface discretizations.

#### 4.1.1 The Nodal Discretization for Vectors

Here the cell discretization of scalar functions and the nodal discretization of vector functions is used. In the interior of the region, the prime operator  $\mathcal{D} = \mathbf{DIV}$ , which is the discrete analog of the divergence, is given by

$$(\mathcal{D}\vec{W})_{(i,j)} = \frac{1}{VC_{(i,j)}} \left\{ \begin{bmatrix} \left(WX_{(i,j)} - WX_{(i+1,j+1)}\right) \left(y_{(i+1,j)} - y_{(i,j+1)}\right) \\ - \left(WX_{(i+1,j)} - WX_{(i,j+1)}\right) \left(y_{(i,j)} - y_{(i+1,j+1)}\right) \right] - \\ \left[ \left(WY_{(i,j)} - WY_{(i+1,j+1)}\right) \left(x_{(i+1,j)} - x_{(i,j+1)}\right) \\ - \left(WY_{(i+1,j)} - WY_{(i,j+1)}\right) \left(x_{(i,j)} - x_{(i+1,j+1)}\right) \right] \right\},$$

$$(4.39)$$

while on the boundary  $\mathcal{D}$  gives an approximation of the normal component of vector. For example on the "bottom boundary" where j=1 and  $i=1,\cdots,M-1$ , a unit normal vector is

$$\left(\frac{y_{(i+1,1)} - y_{(i,1)}}{l\xi_{(i,1)}}, -\frac{x_{(i+1,1)} - x_{(i,1)}}{l\xi_{(i,1)}}\right), \tag{4.40}$$

where, because the problem is two dimensional,  $l\xi = S\eta$  and  $l\eta = S\xi$  are the lengths of the edges of the cell, and then  $\mathcal{D}$  is given by

$$(\mathcal{D}\vec{W})_{(i,0)} = -\left(\frac{WX_{(i,1)} + WX_{(i+1,1)}}{2} \frac{y_{(i+1,1)} - y_{(i,1)}}{l\xi_{(i,1)}} - \frac{WY_{(i,1)} + WY_{(i+1,1)}}{2} \frac{x_{(i+1,1)} - x_{(i,1)}}{l\xi_{(i,1)}}\right). \tag{4.41}$$

## 4.1.2 The Surface Discretization for Vectors

For a cell discretization for scalar functions and a surface discretization for vector functions, the prime operator (discrete divergence  $= \mathcal{D} = \mathbf{DIV}$ ) is defined in the interior of the region by

$$(\mathcal{D}\vec{W})_{(i,j)} = \frac{1}{VC_{(i,j)}} \left\{ (4.42) \left( WS\xi_{(i+1,j)} S\xi_{(i+1,j)} - WS\xi_{(i,j)} S\xi_{(i,j)} \right) + \left( WS\eta_{(i,j+1)} S\eta_{(i,j+1)} - WS\eta_{(i,j)} S\eta_{(i,j)} \right) \right\},$$

while on the boundary  $\mathcal{D}$  gives an approximation of the normal component of a vector:

$$(\mathcal{D} \vec{W})_{(i,0)} = -WS\eta_{(i,1)}, \quad i = 1, \dots, M-1,$$

$$(\mathcal{D} \vec{W})_{(i,N)} = +WS\eta_{(i,N)}, \quad i = 1, \dots, M-1,$$

$$(\mathcal{D} \vec{W})_{(0,j)} = -WS\xi_{(1,j)}, \quad j = 1, \dots, N-1,$$

$$(\mathcal{D} \vec{W})_{(M,j)} = +WS\xi_{(M,j)}, \quad j = 1, \dots, N-1.$$

$$(4.43)$$

# 4.2 The Derived Operator

The derived operator  $\mathcal{G}$  is the discrete analog of the flux operator, and is defined by  $\mathcal{G} = \mathcal{D}^*$ . Here the adjoint is taken in the natural inner product. For the surface discretization on arbitrary grids, it is not possible to write a explicit formula for the components of the operator  $\mathcal{G}$ . However, it is possible to express  $\mathcal{G}$  in terms of  $\mathcal{M}$ ,  $\mathcal{S}$ , and  $\mathcal{D}$ . For the nodal discretization,  $\mathcal{G}$  can be expressed explicitly in terms of  $\mathcal{M}$ ,  $\mathcal{N}$ , and  $\mathcal{D}$ .

#### 4.2.1 The Surface Discretization for Vectors

For the cell discretization of scalar functions and the surface discretization for vectors,  $\mathcal{G}$  is computed by finding the adjoint of  $\mathcal{D}: \mathcal{HS} \to HC$ :

$$(\mathcal{D}\,\vec{W},U)_{HC} = (\vec{W},\mathcal{D}^*U)_{\mathcal{HS}}\,,\tag{4.44}$$

which can be rewritten in terms of the formal inner products as

$$[\mathcal{D}\vec{W}, \mathcal{M}U]_{HC} = [\vec{W}, \mathcal{S}\mathcal{D}^*U]_{\mathcal{HS}}. \tag{4.45}$$

The formal adjoint  $\mathcal{D}^{\dagger}$  of  $\mathcal{D}$  is defined to be the adjoint in the formal inner product, so

$$[\vec{W}, \mathcal{D}^{\dagger} \mathcal{M} U]_{\mathcal{HS}} = [\vec{W}, \mathcal{S} \mathcal{D}^* U]_{\mathcal{HS}}. \tag{4.46}$$

This relationship must be true for all  $\vec{W}$  and U, so

$$\mathcal{D}^{\dagger} \mathcal{M} = \mathcal{S} \mathcal{D}^* \,, \tag{4.47}$$

which gives

$$\mathcal{G} = \mathcal{D}^* = \mathcal{S}^{-1} \, \mathcal{D}^{\dagger} \, \mathcal{M} \,. \tag{4.48}$$

Because S is banded,  $S^{-1}$  is likely to be full (unless S is diagonal). Hence, G is full and has a *non-local* stencil. This is not a serious problem, because we do not need to explicitly form G. The discrete fluxes are

$$\vec{W} = \mathcal{G} U = \mathcal{S}^{-1} \mathcal{D}^{\dagger} \mathcal{M} U, \qquad (4.49)$$

and if the operator S is applied to both sides of this equation, then

$$S\vec{W} = \mathcal{D}^{\dagger} \mathcal{M} U. \tag{4.50}$$

The operators on both sides of this equation have a local stencils.

These equations are similar to the finite element and compact finite difference methods that can be expressed in the form (4.50) with local stencils. (see, for example, [10, 1, 11]).

To find the fluxes for given temperature, from (4.50) we must solve a system of linear equations. The discrete operator S is symmetric positive-definite and with five non-zero elements in each row (see (3.36) and Figure 3.9). In Section 4.6, we discuss possible solution approaches.

The relationship  $(\mathcal{D} \vec{W}, U)_{HC} = (\vec{W}, \mathcal{D}^* U)_{\mathcal{HS}}$  implies that

$$[\vec{W}, \mathcal{D}^{\dagger} \mathcal{M} U]_{\mathcal{HS}} = [\mathcal{D} \vec{W}, \mathcal{M} U]_{HS}. \tag{4.51}$$

The right-hand side of this formula can be evaluated using (4.42) for  $\mathcal{D}$  and summation by parts to give:

$$-\left(\mathcal{D}^{\dagger} \mathcal{M} U\right)_{(i,j)} = \begin{pmatrix} S\xi_{(i,j)} \left(U_{(i,j)} - U_{(i-1,j)}\right) \\ S\eta_{(i,j)} \left(U_{(i,j)} - U_{(i,j-1)}\right) \end{pmatrix}. \tag{4.52}$$

#### 4.2.2 The Nodal Discretization for Vectors

In the case of the cell discretization of scalar functions and the nodal discretization of vectors, an argument similar to that given in the previous section gives

$$\mathcal{D}^{\dagger} \mathcal{M} = \mathcal{N} \mathcal{D}^*, \tag{4.53}$$

which gives

$$\mathcal{G} = \mathcal{D}^* = \mathcal{N}^{-1} \, \mathcal{D}^{\dagger} \, \mathcal{M} \,. \tag{4.54}$$

Note that the operator  $\mathcal{D}$  here is not the same as in the previous section. As before, the fluxes are given by

$$\vec{W} = \mathcal{G} U = \mathcal{N}^{-1} \mathcal{D}^{\dagger} \mathcal{M} U, \qquad (4.55)$$

where  $\vec{W} = (WX, WY)$ . Applying the operator  $\mathcal{N}$  to both sides gives

$$\mathcal{N} \vec{W} = \mathcal{D}^{\dagger} \mathcal{M} U. \tag{4.56}$$

To simplify the notation, introduce  $\vec{F}$ :

$$\mathcal{N} \vec{W} = \vec{F} = \mathcal{D}^{\dagger} \mathcal{M} U. \tag{4.57}$$

where  $\vec{F} = (FX, FY)$ . As before, the operator  $\mathcal{N}$  is given by a two by two block of operators, but in this case, the blocks are diagonal, so  $\mathcal{N}$  can be inverted explicitly by solving the left-hand equation in (4.57). The explicit form of this equations is

$$n11_{(i,j)} WX_{(i,j)} + n12_{(i,j)} WY_{(i,j)} = FX_{(i,j)},$$

$$n12_{(i,j)} WX_{(i,j)} + n22_{(i,j)} WY_{(i,j)} = FY_{(i,j)},$$

$$(4.58)$$

and the solution is

$$WX_{(i,j)} = \frac{FX_{(i,j)} n22_{(i,j)} - FY_{(i,j)} n12_{(i,j)}}{n11_{(i,j)} n22_{(i,j)} - n12_{(i,j)}^{2}},$$

$$WY_{(i,j)} = -\frac{FX_{(i,j)} n12_{(i,j)} - FY_{(i,j)} n11_{(i,j)}}{n11_{(i,j)} n22_{(i,j)} - n12_{(i,j)}^{2}}.$$
(4.59)

The relationship  $(\mathcal{D}\vec{W}, U)_{HC} = (\vec{W}, \mathcal{D}^* U)_{\mathcal{HN}}$  implies that

$$[\vec{W}, \mathcal{D}^{\dagger} \mathcal{M} U]_{\mathcal{H}\mathcal{N}} = [\mathcal{D} \vec{W}, \mathcal{M} U]_{HC}. \tag{4.60}$$

The right-hand side of the last formula can be evaluated using Formula (4.39) for  $\mathcal{D}$  and summation by parts to give an explicit formula for  $\vec{F} = \mathcal{D}^{\dagger} \mathcal{M} U$ 

$$\begin{array}{lcl} -FX_{(i,j)} & = & 0.5 \, \left\{ \left( y_{(i,j+1)} - y_{(i+1,j)} \right) \, U_{(i,j)} + \left( y_{(i-1,j)} - y_{(i,j+1)} \right) \, U_{(i-1,j)} \right. \\ & + & \left. \left( y_{(i,j-1)} - y_{(i-1,j)} \right) \, U_{(i-1,j-1)} + \left( y_{(i+1,j)} - y_{(i,j-1)} \right) \, U_{(i,j-1)} \right\} \, . \end{array}$$

The formula for FY is given by changing y to -x in the previous.

# 4.3 The Discrete Operator Equations

For both discretizations, the discrete analog of the continuum operator  $\Omega$  (2.11), is defined by

$$(\Omega U)_{(i,j)} = \begin{cases} 0, & \text{in the interior,} \\ \alpha_{(i,j)} U_{(i,j)}, & \text{on the boundary.} \end{cases}$$
(4.61)

The finite difference method approximating the first-order system (2.6), written as an analog of the continuum-operator system (2.13), is

$$\Omega U + \mathcal{D}\vec{W} = F , \quad \vec{W} = \mathcal{G} U . \tag{4.62}$$

Then the discretization of the second-order equation (1.1), which is an analog of the operator equation (2.8), is

$$A U = (\Omega + \mathcal{D} \mathcal{G}) U = F.$$
 (4.63)

For both discretizations, in the interior of the cells, Equation (4.63) is

$$\left(\mathcal{D}\vec{W}\right)_{(i,j)} = \mathbf{DIV}\,\vec{W}_{(i,j)} = f_{(i,j)}. \tag{4.64}$$

The approximation of the boundary conditions is

$$\left(\mathcal{D}\vec{W}\right)_{(i,j)} + \alpha_{(i,j)} \, U_{(i,j)} = \psi_{(i,j)} \,, \tag{4.65}$$

where, on the boundary, the operator  $\mathcal{D}$  is an approximation of the normal component of the vector.

Also, for both discretizations, the fluxes are determined from

$$\vec{W} = \mathcal{G} u. \tag{4.66}$$

For the cell-node discretization, the operator  $\mathcal{G}$  can be constructed from (4.59) and (4.61) and there are local explicit formulas for the fluxes. (see Figure 4.10). For the cell-surface discretization the operator  $\mathcal{G}$  is non-local (see Section 4.2.1) and, consequently, there is no local explicit equation for the fluxes.

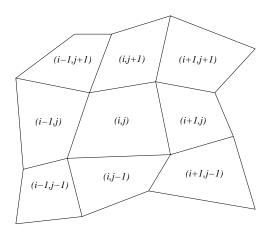


Figure 4.10: The nine-cell stencil for the cell-node discretization for the operator A

# 4.4 The Discrete Operators on a Rectangular Grid

On orthogonal grids and for diagonal K, the interior discretizations discussed in this paper reduce to well-known discretizations. In this section we consider the case where K = kI on a rectangular grid with the cell sides  $S\xi_{(i,j)} = hY$  and  $S\eta_{(i,j)} = hX$  and cell volume  $VC_{(i,j)} = hX hY$ 

#### 4.4.1 The Cell-Node Discretization

For the cell-node discretization, the DIV operator (4.39) is

$$(\mathbf{DIV}\,\vec{W})_{(i,j)} = \frac{0.5\left(WX_{(i+1,j)} + WX_{(i+1,j+1)}\right) - 0.5\left(WX_{(i,j)} + WX_{(i,j+1)}\right)}{hX} + \frac{0.5\left(WY_{(i+1,j+1)} + WY_{(i,j+1)}\right) - 0.5\left(WY_{(i+1,j)} + WY_{(i,j)}\right)}{hY},$$
(4.67)

which is a natural discretization for a rectangular grid.

The operator  $\mathcal{G}$  from (4.56) is more complicated. First, the operator  $\mathcal{N}$  is given by (3.29) through (3.32). On a rectangular grid the formula for n11 becomes

$$n11_{(i,j)} = 0.25 hX hY \sum_{k,l=0}^{1} \frac{1}{\mathbf{k}_{(i-k,j-l)}}.$$
 (4.68)

This can be written in terms of the two-dimensional cell-to-node harmonic average

$$\tilde{\mathbf{k}}_{(i,j)} = \left(\frac{1}{4} \sum_{k,l=0}^{1} \frac{1}{k_{(i-k,j-l)}}\right)^{-1} . \tag{4.69}$$

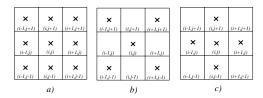


Figure 4.11: The Stencil of the Laplacian for: a) Cell-Node Discretization: Rectangular Grid b) Cell-Node Discretization: Square Grid c) Cell-Surface Discretization: Rectangular Grid

Next, (4.61) is

$$-FX_{(i,j)} = 0.5 hY \left\{ U_{(i,j)} - U_{(i-1,j)} - U_{(i-1,j-1)} + U_{(i,j-1)} \right\}$$

$$= hY \left\{ 0.5 \left( U_{(i,j)} + U_{(i,j-1)} \right) - 0.5 \left( U_{(i-1,j)} + U_{(i-1,j-1)} \right) \right\},$$

$$(4.70)$$

which gives

$$WX_{(i,j)} = \tilde{\mathbf{k}}_{(i,j)} \left\{ \frac{0.5 \left( U_{(i,j)} + U_{(i,j-1)} \right) - 0.5 \left( U_{(i-1,j)} + U_{(i-1,j-1)} \right)}{hX} \right\},$$

$$WY_{(i,j)} = \tilde{\mathbf{k}}_{(i,j)} \left\{ \frac{0.5 \left( U_{(i,j)} + U_{(i-1,j)} \right) - 0.5 \left( U_{(i,j-1)} + U_{(i-1,j-1)} \right)}{hY} \right\}. \tag{4.71}$$

The expressions in curly braces is an approximation for the derivatives  $\partial u/\partial x$  and  $\partial u/\partial y$ . The stencil for the discrete analog of the Laplacian **div grad** is given by choosing  $\mathbf{K} = \mathbf{I}$ :

$$\frac{1}{4hX^{2}} + \frac{1}{4hY^{2}} \left| \frac{-1}{2hX^{2}} + \frac{1}{2hY^{2}} \right| \frac{1}{4hX^{2}} + \frac{1}{4hY^{2}}$$

$$\frac{1}{2hX^{2}} + \frac{-1}{2hY^{2}} \left| \frac{-1}{hX^{2}} + \frac{-1}{hY^{2}} \right| \frac{1}{2hX^{2}} + \frac{-1}{2hY^{2}}$$

$$\frac{1}{4hX^{2}} + \frac{1}{4hY^{2}} \left| \frac{-1}{2hX^{2}} + \frac{1}{2hY^{2}} \right| \frac{1}{4hX^{2}} + \frac{1}{4hY^{2}}$$

$$(4.72)$$

When the grid is square, the stencil becomes the five-point stencil shown in Figure 4.11. On a square grid, the cell-node Laplacian has a nontrivial null space that includes the checker board mode.

#### 4.4.2 The Cell-Surface Discretization

For the surface discretization of vectors, Formula (4.42) gives the operator DIV as

$$(\mathbf{DIV}\,\vec{W})_{(i,j)} = \frac{WS\xi_{(i+1,j)} - WS\xi_{(i,j)}}{hX} + \frac{WS\eta_{(i,j+1)} - WS\eta_{(i,j)}}{hY},\tag{4.73}$$

which is also a natural discretization for rectangular grid.

The operator  $\mathcal{G}$  is computed from Formula (4.50), and involves the operator  $\mathcal{S}$  given by (3.36). For orthogonal grids  $\mathcal{S}$  is diagonal (the sines of all angles are one and the cosines of all angles are zero). Also, all  $V_{(i,j)}^{(p,q)}$  are equal to 0.25 hX hY. Using (4.52), Equation (4.50) for internal cells becomes

$$\left(0.5 hX hY \sum_{k=0,1} \frac{1}{\mathbf{k}_{(i-k,j)}}\right) WS\xi_{(i,j)} = hY \left(U_{(i,j)} - U_{(i-1,j)}\right)$$
(4.74)

or

$$WS\xi_{(i,j)} = \frac{2 k_{(i-1,j)} k_{(i,j)}}{k_{(i-1,j)} + k_{(i,j)}} \frac{U_{(i,j)} - U_{(i-1,j)}}{hX}.$$
(4.75)

Thus, on rectangular grids, the cell-surface discretization leads to the well-known harmonic average for the coefficient k in the  $\xi$  direction. The formula for  $WS\eta$  is similar:

$$WS\eta_{(i,j)} = \frac{2 \,\mathbf{k}_{(i,j-1)} \,\mathbf{k}_{(i,j)}}{\mathbf{k}_{(i,j-1)} + \mathbf{k}_{(i,j)}} \frac{U_{(i,j)} - U_{(i,j-1)}}{hY}, \tag{4.76}$$

and contains the harmonic average for k in  $\eta$  direction.

The fluxes on the boundary are one-sided differences. For example, on the left boundary

$$WS\xi_{(1,j)} = k_{(1,j)} \frac{U_{(1,j)} - U_{(0,j)}}{0.5 hX}, \quad WS\eta_{(i,1)} = k_{(i,1)} \frac{U_{(i,1)} - U_{(i,0)}}{0.5 hY}.$$
(4.77)

The discrete analog of the Laplacian div grad is

$$\frac{U_{(i+1,j)} - 2U_{(i,j)} + U_{(i-1,j)}}{hX^2} + \frac{U_{(i,j+1)} - 2U_{(i,j)} + U_{(i,j-1)}}{hY^2}$$
(4.78)

which is the usual five point approximation on a rectangular grid with the stencil shown in Figure 4.11.

# 4.5 Theoretical Properties of the Algorithms

For the cell-surface discretization, the properties for the operators  $\operatorname{div}$  and  $\operatorname{grad}$  were investigated in Shashkov and Steinberg [15], where it was shown that the divergence of a constant vector is zero, that for smooth grids the point truncation errors for the divergence  $\operatorname{DIV}$  and for the gradient  $\operatorname{GRAD}$  are second order, and for general grids,  $\operatorname{DIV}$  and  $\operatorname{GRAD}$  are first-order accurate, and that the  $\operatorname{DIV}$  is exact for integral truncation error. In the Appendix A, we show, using a rather lengthy geometric calculation, that for piecewise constant  $\operatorname{K}$ , the discrete analog of  $\operatorname{K}\operatorname{grad}$  is exact on piecewise linear functions. In the following subsection, it is shown that, for the cell-surface discretization, the null space of  $\operatorname{G}$  is exactly the constants, while for nodal discretization on square grids, the null space of  $\operatorname{G}$  contains highly oscillatory checker board grid functions in addition to the constants.

In Section 5, the approximation properties of the variable-coefficient Laplacian div K grad are numerically shown to confirm these theoretical results.

## 4.5.1 The Null Space of the Operator $\mathcal{G}$

For the cell-surface discretization, we prove that  $\mathcal{G}U$  is zero if and only if U is constant. Formula (4.48) gives

$$\mathcal{G} U = \mathcal{S}^{-1} \mathcal{D}^{\dagger} \mathcal{M} U, \qquad (4.79)$$

and then, if U is a constant, (4.52) shows that  $\mathcal{D}^{\dagger} \mathcal{M} U = 0$ , so  $\mathcal{G} U = 0$ .

Conversely, assume that  $\mathcal{G}U=0$ . Formula (4.79) and the fact that operator  $\mathcal{S}$  is positive definite gives

$$\mathcal{D}^{\dagger} \mathcal{M} U = 0. \tag{4.80}$$

This and Formula (4.52) then gives:

$$egin{aligned} U_{(i,j)} - U_{(i-1,j)} &= 0 \; ; \quad i = 1, \ldots, M \; ; \quad j = 1, \ldots, N-1 \; ; \ U_{(i,j)} - U_{(i,j-1)} &= 0 \; ; \quad i = 1, \ldots, M-1 \; ; \quad j = 1, \ldots, N \; ; \end{aligned}$$

which implies that U is a constant. Therefore the null space of the discrete operator  $\mathcal{G}$  is the constant functions, exactly as for the differential operator k grad.

For the cell-node discretization, the situation is quite different: in the case of a square grid, Formula (4.61) shows that both the constant function  $U_{(i,j)} = 1$  and the "checkerboard" function  $U_{(i,j)} = (-1)^{i+j}$  are satisfy  $\mathcal{G}U = 0$ , as do any linear combinations of these functions. The checkerboard solution or mode is well known, especially in computational Lagrangian gas dynamics, where it leads to the so-called "hour-glassing" instability [13]. In case of elliptic equations, this mode leads to the presence of high-frequency noise in the solution, as illustrated by the numerical examples in Section 5.

# 4.6 Solving the System of Linear Equations

The discrete equations for both the cell-node and the cell-surface discretization have the form (4.62):

$$\Omega U + \mathcal{D} \vec{W} = F, \quad \vec{W} = \mathcal{G} U. \tag{4.81}$$

The fluxes can be eliminated from this system to obtain an equation for U:

$$A U = \Omega U + \mathcal{D} \mathcal{G} U = F, \qquad (4.82)$$

where A is symmetric and positive definite.

For the cell-node discretization, the operator  $\mathcal{A}$  is symmetric and positive definite, and has a local 9 cell stencil. In our numerical examples to compare the accuracy of the methods, we used a simple SOR iteration to solve this system. More efficient iteration methods, such as multigrid [3, 4, 19, 2] or incomplete Cholesky conjugate gradient methods [8], could also have been used to solve these equations.

For the case of cell-surface discretization, the operators  $\mathcal{G}$  and  $\mathcal{A}$  are non-local and, therefore, algorithms that require explicit expressions for  $\mathcal{A}$  are impractical for large problems. The equations can be formulated so that algorithms, such as preconditioned conjugate gradient methods, requiring only a multiplication of a vector by  $\mathcal{A}$  can be used. Given U,

 $\mathcal{A}\,U$  can be computed efficiently by solving (4.50)  $\mathcal{S}\,\vec{W}=\mathcal{D}^\dagger\,\mathcal{M}\,U$ , for  $\vec{W}$  and evaluating  $\mathcal{A}\,U=\Omega\,U+\mathcal{D}\vec{W}$ . All operators in this formulation are explicitly known and local. Moreover, because  $\mathcal{S}$  is a positive definite symmetric local operator, the equation for  $\vec{W}$  can be solved efficiently with iterative methods. On orthogonal grids  $\mathcal{S}$  is a diagonal operator, and all steps of this procedure are local.

Other efficient algorithms to solve this system include the family of two-level gradient methods, including the minimal residual method, the minimal correction method, and the minimal error method. All these methods can be written as

$$B\frac{U^{(s+1)} - U^{(s)}}{\tau_s} + A U^{(s)} = F, \qquad (4.83)$$

where  $U^{(s)}$  is approximate solution to  $U^{n+1}$  on iteration number s,  $\tau_s$  some iteration parameter, and operator B is preconditioner. A family of three-level iteration methods, which require only the computation of AU include the three-level conjugate-direction methods, like the conjugate gradient method. All these methods can be written as

$$B U^{(s+1)} = \alpha_{s+1} (B - \tau_{s+1} A) U^{(s)} + (1 - \alpha_{s+1}) B U^{(s-1)} + \alpha_{s+1} \tau_{s+1} F,$$
  
 $B U^{(1)} = (B - \tau_1 A) U^{(0)} + \tau_1 F.$ 

The effectiveness of these methods strongly depends on the choice of a preconditioner. The simplest Jacobi type preconditioner approximates  $\mathcal{S}$  by its diagonal blocks. This is exact for orthogonal grids and produces a five-cell symmetric positive-definite operator corresponding to removing the mixed derivatives from the variable-coefficient Laplacian on non-orthogonal grids.

### 4.6.1 Solving for the Fluxes

Given U, the system (4.50) must be solved to obtain the flux. In our examples, we used the block Gauss-Seidel algorithm:

$$S_{11} W S \xi^{(s+1)} + S_{12} W S \eta^{(s)} = \mathcal{F} \mathcal{X},$$
 (4.84)

$$S_{21} W S \xi^{(s+1)} + S_{22} W S \eta^{(s+1)} = \mathcal{F} \mathcal{Y},$$
 (4.85)

where (s) and (s+1) are iteration indices. Equation 4.84 gives all  $WS\xi^{(s+1)}$  fluxes for the new iteration (s+1), and then from equation 4.85, all fluxes  $WS\eta^{(s+1)}$  can be found. Because the matrix of the operators  $S_{11}$  and  $S_{22}$  are diagonal, systems 4.84 and 4.85 can be solved explicitly. Because the operators for these equations are symmetric and positive definite, the block Gauss-Seidel method always converges.

Another approach to solving (4.50) is to use the fact that the operators  $S_{11}$ , and  $S_{22}$  are diagonal, and then eliminate either  $WS\xi$  or  $WS\eta$  from 4.84. For example, to the equation resulting from eliminating  $WS\eta$  is

$$\left(S_{11} + S_{12} S_{22}^{-1} S_{21}\right) W S \xi = \mathcal{F} \mathcal{X} - S_{22} \mathcal{F} \mathcal{Y}.$$
 (4.86)

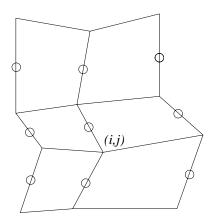


Figure 4.12: The Stencil for  $WS\xi$ 

The operator of this system,

$$\left(S_{11} + S_{12} S_{22}^{-1} S_{21}\right)$$
, (4.87)

is symmetric and positive definite, and has the local stencil shown in 4.12.

# 5 Numerical Examples

The examples in this section, summarized in Table 5.1, solve the elliptic PDE (1.1) and were chosen to illustrate how the cell-node and cell-surface algorithms perform for discontinuous nondiagonal K and on nonuniform grids. The first example verifies that the algorithms are exact for linear solutions on random logically-rectangular grids. Examples 2, 3, and 4 are used to compare the convergence rates of the algorithms for various kinds of coefficients and grids. Examples 5 and 6 show that the algorithms can compute accurate total fluxes for very rough K. Example 7 confirm that the algorithms produce reasonable results for the case of non-diagonal and discontinuous K and non-uniform grids, while Example 8 verify that for non-diagonal and discontinuous K and non-uniform grids, but where the solution is piecewise linear, that the cell-surface algorithm is exact and the cell-node algorithm is second-order accurate.

The asymptotic truncation error  $E_h$  on a grid of  $M \times N$  nodes,

$$h = \max\{\frac{1}{M-1}, \frac{1}{N-1}\},\tag{5.1}$$

is estimated by

$$||E_h|| = C h^q + O(h^{q+1}), (5.2)$$

where q is the order of the error, and the constant C, the convergence-rate constant, is independent of h, and  $||\cdot||$  is some norm.

		K	K	grid
#	name	diag	cont	uniform
5.1		yes	no	no
5.2	MacKinnon	yes	no	no
5.3	Crumption 1	no	yes	yes
5.4	Crumption 2	no	no	yes
5.5	Jikov	yes	no	yes
5.6	Durlofsky 1	yes	no	yes
5.7	Durlofsky 2	no	no	no
5.8	Das	no	no	no

Table 5.1: Summary of Examples

In the numerical examples the truncation errors were evaluated on a sequence of grids  $h, h/2, h/4, \ldots$  and the convergence rates estimated from the ratio between the norms of the errors (5.2) and

$$||E_{h/2}|| = C \frac{h^q}{2^q} + O(h^{q+1}). {(5.3)}$$

The order of convergence q can be estimated by

$$q pprox \log_2 \frac{||E_h||}{||E_{h/2}||}$$
 (5.4)

In the numerical experiments, continuum functions are discretized using the projection operator

$$(p_h\,u)_{i,j} = u(x_{i,j}^c\,,y_{i,j}^c)$$

where  $x_{i,j}^c$ ,  $y_{i,j}^c$  are the coordinates of geometric center of the cell.

The convergence rates were estimated using both the maximum norm

$$E_{max} = \parallel U - p_h u \parallel_{max} = \max_{i,j} |U_{i,j} - (p_h|u)_{i,j}|$$

and the mean-square norm

$$E_{L_{2}} = \parallel U - p_{h}u \parallel_{L_{2}} = \left( \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} (U_{i,j} - (p_{h}u)_{i,j})^{2} VC_{i,j} \right)^{\frac{1}{2}},$$
 (5.5)

where U is the solution of the finite-difference method and u is the exact solution.

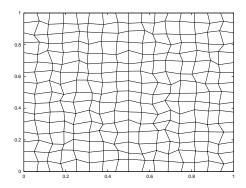


Figure 5.13: Example 5.1 and 5.2: A Random Grid, M = N = 17

## 5.1 Piecewise Linear Solutions

The first test problem has a diagonal diffusion coefficient, that is,  $\mathbf{K}(x,y) = \mathbf{k}(x,y)\mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix, and  $\mathbf{k}$  is piecewise constant:

$$\mathbf{k} = \begin{cases} \mathbf{k}_1, \ 0 < x < 0.5 \\ \mathbf{k}_2, \ 0.5 < x < 1 \end{cases}$$
 (5.6)

On the top and bottom boundary, the flux is equal to zero, while on the left and right boundaries, the mixed boundary conditions

$$u - 2 k_1 \frac{\partial u}{\partial x} = 0, \quad u + 2 k_2 \frac{\partial u}{\partial x} = 1,$$
 (5.7)

are used. The steady-state solution of this problem is

$$u = \begin{cases} \frac{k_2 x + 2 k_1 k_2}{0.5 (k_1 + k_2) + 4 k_1 k_2}, & 0 < x < 0.5, \\ \frac{k_1 x + 2 k_1 k_2 + 0.5 (k_2 - k_1)}{0.5 (k_1 + k_2) + 4 k_1 k_2}, & 0.5 < x < 1, \end{cases}$$
(5.8)

that is, the solution is a piecewise linear function.

This 1-D problem is solved on the 2-D random grid in the unit square shown in Figure 5.13. This mesh was generated by perturbing an initial uniform cell width of w by displacing each node to a random position on a circle of radius 0.2w centered about the original position of the node. Because our methods require that any discontinuity line coincide with a grid line, the line x = 0.5 is kept as a grid line, so only the y coordinates for the nodes are changed. Note that such grids satisfy the regularity condition for grids formulated in Section 3. This example verifies that our cell-surface method is exact for this problem. The isolines of the approximate solution are shown in Figure 5.14.

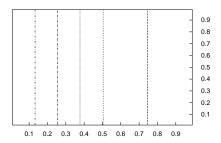


Figure 5.14: Example 5.1: The Isolines for the Approximate Solution, M=N=17

## 5.2 Piecewise Quadratic Solutions

The second test problem is from *MacKinnon and Carey* [12], where the diffusion equation has a constant right-hand side of one, and the diffusion coefficient is the same as in the previous problem. The exact solution is one dimensional:

$$u(x) = \begin{cases} a_1 \frac{x^2}{2} + b_1 x, & 0 \le x \le \frac{1}{2}, \\ a_2 \frac{x^2}{2} + b_2 x + c_2, & \frac{1}{2} \le x \le 1, \end{cases}$$
(5.9)

where

$$a_i = rac{-1}{\mathrm{k}_i} \,, \quad b_1 = -0.25 \, (3 \, a_2 + a_1) \, rac{\mathrm{k}_2}{\mathrm{k}_1 + \mathrm{k}_2} \,, \quad b_2 = rac{\mathrm{k}_2}{\mathrm{k}_1} \, b_1 \,, \quad c_2 = -(b_2 + 0.5 \, a_2) \,. \quad (5.10)$$

For this problem, the flux is equal to zero on the top and bottom boundaries and for the other boundaries, the following are used: on the left boundary

$$u - 2 \mathbf{k}_1 \frac{\partial u}{\partial x} = -2 \mathbf{k}_1 b_1, \qquad (5.11)$$

and on the right boundary

$$u + 2 k_2 \frac{\partial u}{\partial x} = 0.5 a_2 + b_2 + c_2 + 2 k_1 (a_2 + b_2).$$
 (5.12)

As in the previous problem, the 1-D problem is solved on a 2-D random grid on the unit square that is shown in Figure 5.13. The approximate solution for M=N=41 is shown in Figure 5.15. Results of the convergence tests on random grids are presented in Table 5.2, where the first column gives the number of grids points with N=M, the next two columns give the maximum and mean-square error, and the final two columns give the estimated convergence rate as second-order in both the max and mean-square norms.

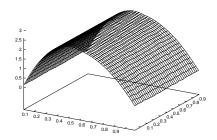


Figure 5.15: Example 5.2: The Piecewise Quadratic Solution

M-1	max	$L_2$ -norm	$q_{max}$	$oldsymbol{q}_2$
	norm			
10	5.98E-2	3.40E-2	1.88	1.99
20	1.62E-2	8.54E-3	1.98	1.96
40	4.10E-3	2.18E-3	-	ı

Table 5.2: Example 5.2: Convergence Rates for a Random Grid: MacKinnon and Carey Test Problem

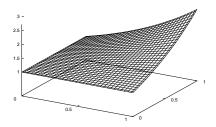


Figure 5.16: Example 5.3: The Pressure for the Cell-Node Discretization, M=N=33

Method	M	max	$L_2$ -norm	$q_{max}$	$q_2$
		norm			
Cell-Node	17	3.74E-3	1.06E-3	1.84	2.03
	33	9.66E-4	2.58E-4	1.97	2.02
	65	2.45E-4	$6.36  ext{E-5}$	-	-
Cell-Surface	17	5.11E-3	1.68E-3	1.92	2.01
	33	1.35E-3	4.15E-4	1.95	2.09
	65	3.48E-4	9.73E-5	-	-

Table 5.3: Example 5.3: Convergence Rates: Cell-Node and Cell-Surface Discretizations

# 5.3 Non-Diagonal Continuous K

Problem 1, from Crumpton, Shaw, and Ware [2], has a constant non-diagonal K, defined on the unit square with Dirichlet boundary conditions obtained from the exact solution. The permeability is

$$\mathbf{K} = \left(\begin{array}{cc} 2 & 1\\ 1 & 2 \end{array}\right) \,, \tag{5.13}$$

where **K** is a positive definite matrix. The true solution is  $u = e^{xy}$ , which corresponds to this right-hand side

$$f(x,y) = -2(1+x^2+xy+y^2)e^{xy}. (5.14)$$

Figure 5.16 displays the approximate solution for M=N=33 and the cell-node discretization. Convergence-rate data for the cell-node and cell-surface discretizations are in given in Table 5.3, which indicates second-order convergence rates for both methods, both in max and  $L_2$  norms.

Method	M	max	$L_2$ -norm	$q_{max}$	$q_2$
		norm			
Cell-Node	17	1.00E-2	7.98E-3	0.82	1.17
	33	5.66E-3	3.53E-3	0.90	1.03
	65	3.02E-3	1.72E-3	-	-
Cell-Surface	17	9.63E-3	7.05E-3	1.89	2.02
	33	2.59E-3	1.73E-3	1.96	2.12
	65	6.72E-4	3.96E-4	_	-

Table 5.4: Example 5.4: Convergence Rates: Cell-Node and Cell-Surface Discretization

# 5.4 Non-Diagonal Discontinuous K

This test problem from Crumpton, Shaw, and Ware [2] is defined on the square  $[-1,1] \times [-1,1]$ , with Dirichlet boundary conditions. The diffusion coefficient is given by

$$\mathbf{K} = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & x < 0, \\ \alpha \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, & x > 0, \end{cases}$$
 (5.15)

where the parameter  $\alpha$  is used to vary the strength of the discontinuity at x = 0. The exact solution is:

$$u(x,y) = \begin{cases} [2\sin(y) + \cos(y)] \alpha x + \sin(y), & x < 0, \\ \exp(x)\sin(y), & x > 0, \end{cases}$$
 (5.16)

The right-hand side, which corresponds to this solution, is also discontinuous:

$$f(x,y) = \begin{cases} [-2\sin(y) - \cos(y)] \alpha x - \sin(y), & x < 0, \\ 2\alpha \exp(x) \cos(y), & x > 0, \end{cases}$$
 (5.17)

No special discretization is needed for the right-hand side because our method uses a cell-centered discretization for u and the discrete analog of operator  $\operatorname{div} \mathbf{Kgrad}$  whose domain and range coincide with HC, and the right-hand side is also assumed to be given in the cells and the material discontinuity coincides with a grid line.

Figure 5.17 displays the isolines for the approximate solution for M=N=17, and  $\alpha=1$ . Table 5.4 gives the convergence-rate data for both the cell-node and the cell-surface discretizations. This data verifies the second-order convergence rate for cell-surface algorithm and a first-order convergence rate for cell-node algorithm.

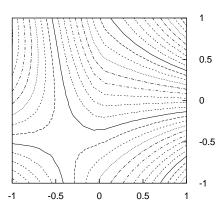


Figure 5.17: Example 5.4: The Isolines for the Pressure, M=N=17

$k_{I}$	$k_2$	$k_I$	$k_2$	$k_{I}$	$k_2$	$ k_1 $	$k_2$
$k_2$	$k_{I}$	$k_2$	$k_{I}$	$k_2$	$k_{I}$	$k_2$	$k_I$
$k_{I}$	$k_2$	$k_{I}$	$k_2$	$k_{I}$	$k_2$	$k_{I}$	$k_2$
$k_2$	$k_{I}$	$k_2$	$k_{I}$	$k_2$	$k_I$	$k_2$	$k_{I}$
$\overline{k_I}$	$k_2$	$k_I$	$k_2$	$k_I$	$k_2$	$k_{I}$	$k_2$

Figure 5.18: Example 5.5: Pattern for Homogenization

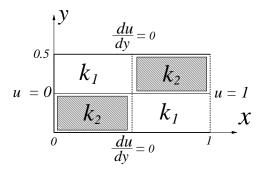


Figure 5.19: Example 5.5: The Homogenization Problem

#### 5.5 Diagonal Discontinuous K: Homogenization Example

The next example, given in Jikov, Kozlov and Oleinik [7], is related to theory of homogenization of differential operators. Assume that a material fills all of space with the pattern which is presented in Figure 5.18. In the shaded areas  $\mathbf{K} = \mathbf{k}_2 \mathbf{I}$ , while in the other areas  $\mathbf{K} = \mathbf{k}_1 \mathbf{I}$ . The homogenization problem is to find some effective matrix  $\tilde{\mathbf{K}}$  that describes the average diffusion coefficient of such a media as the cell size shrinks to zero. In [7] p.37, it is shown that, for such a pattern, the effective matrix is:

$$\tilde{\mathbf{K}} = \sqrt{\mathbf{k}_1 \, \mathbf{k}_2} \, \mathbf{I} \,, \tag{5.18}$$

that is, this media can be described by using a diagonal matrix with a geometrically averaged diffusion coefficient.

To reproduce this result numerically, the diffusion equation was solved on the rectangle  $[0,1] \times [0,.5]$  using a diagonal matrix with  $\mathbf{K} = \mathbf{k}_1 \mathbf{I}$ ,  $\mathbf{k}_1 = 1$  and  $\mathbf{K} = \mathbf{k}_2 \mathbf{I}$ ,  $\mathbf{k}_2 = 100$ , as in Figure 5.19. The boundary conditions are that there are no fluxes through the top and bottom boundaries and Dirichlet conditions on right and left boundary.

According to [7], such a media can be described by an effective diffusion coefficient  $\tilde{\mathbf{K}}$ , and then the flux in x direction through the configuration, which is presented in Figure 5.19, equals

$$FLUX_{x} = -\sqrt{k_{1} k_{2}} \frac{u_{right} - u_{left}}{x_{right} - x_{left}} = -\sqrt{k_{1} k_{2}}.$$
 (5.19)

In this example,  $FLUX_x = 10$ .

The calculations were performed on a square grid, where the grid lines coincide with the discontinuity lines, and the flux through the system was computed as follows

$$(FLUX_x)_{numerical}^{CS} = \frac{\sum_{j=1}^{N-1} WS\xi_{1,j} S\xi_{1,j}}{\sum_{j=1}^{N-1} S\xi_{1,j}}$$

$$= hY \frac{\sum_{j=1}^{N-1} WS\xi_{1,j}}{0.5},$$
(5.20)

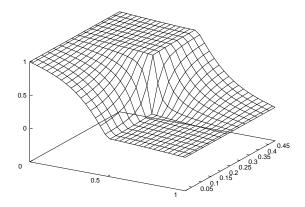
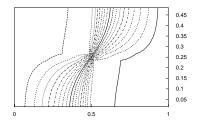
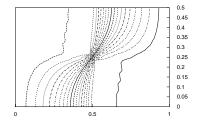


Figure 5.20: Example 5.5: The Pressure for Cell-Surface Discretization





a) Cell-Surface Discretization

b) Cell-Node Discretization

Figure 5.21: Example: 5.5 The Isolines for the Pressure for the Cell-Surface and Cell-Node Discretization

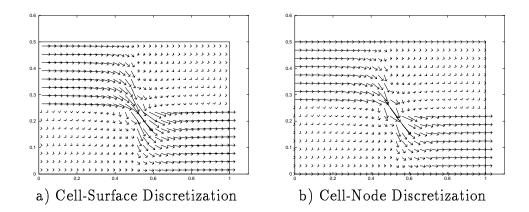


Figure 5.22: Example:5.5 Velocity field for Cell-Node and Cell-Surface Discretization

$\operatorname{Method}$	Μ	approx.	exact	error	q
		flux	flux		
Cell-Node	17	6.8	10	3.2	0.71
	33	8.04	10	1.96	1.22
	65	9.16	10	0.84	-
Cell-Surface	17	7.24	10	2.76	0.83
	33	8.45	10	1.55	1.78
	65	9.55	10	0.45	-

 $\label{local-convergence} \begin{tabular}{lll} Table 5.5: $Example: 5.5: Convergence Rates for Flux: Cell-Node and Cell-Surface Discretization \\ \end{tabular}$ 

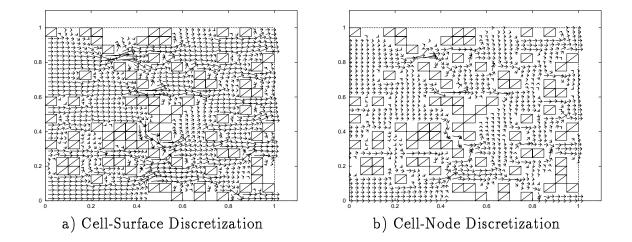


Figure 5.23: Example 5.6: Velocity Field for Sand-Shale Problem

$$(FLUX_x)_{numerical}^{CN} = rac{\sum_{j=1}^{N-1} 0.5 (WX_{1,j} + WX_{1,j+1}) S\xi_{1,j}}{\sum_{j=1}^{N-1} S\xi_{1,j}}$$
  
=  $hY \sum_{j=1}^{N-1} (WX_{1,j} + WX_{1,j+1}),$ 

where hY = 1/(N-1).

Figure 5.20 displays the pressure for the cell-surface discretization while Figures 5.21 and 5.22 show the isolines of pressure and velocity fields for both discretizations. Note that the cell-surface discretization profiles are much smoother than the cell-node solution.

Table 5.5 contains the convergence-rate data for the total flux for both types of discretizations. This table shows that the convergence rate for cell-node discretization is approximately first-order, and that for the cell-surface discretization the convergence rate is approximately second-order.

### 5.6 Flow Through a Sand-Shale System

This example is from Durlofsky [6] and is defined on the unit square, with the boundary conditions: u = 0 along x = 0, u = 1 along x = 1, and the flux is equal to zero along y = 0 and y = 1. The permeability field is generated by randomly placing shale blocks, of total area fraction 0.2, throughout the sand on regular grid of dimension  $20 \times 20$  (see Figure 5.23). In the example, the permeabilities of both the sand  $(k_{sand})$  and shale  $(k_{shale})$  are taken to be uniform and isotropic,  $\mathbf{K} = k \mathbf{I}$ , with  $k_{sand} = 1$  and  $k_{shale} = 10^{-6}$ . The "exact" flux through the system (which is flux obtained on very fine grid) is 0.5202 (see [6]).

Figure 5.23 shows the velocity fields for both types of discretization. The streamlines for the case of the cell-surface discretization, which were obtained from the vector field by

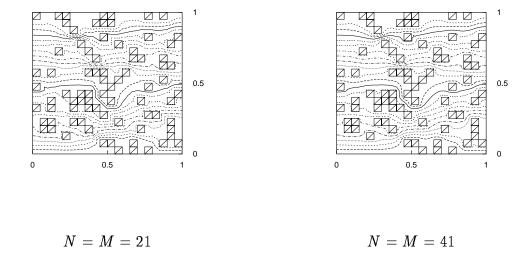
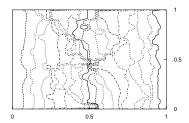


Figure 5.24: Example 5.6: Streamlines for Sand-Shale Problem, M=N=21 and M=N=41

Method	M	approx.	exact	error	q
		flux	flux		
Cell-Node	21	0.022	0.5205	0.498	1.39
	41	0.33	0.5205	0.190	1.07
	81	0.43	0.5205	0.090	-
Cell-Surface	21	0.45	0.5205	0.070	1.22
	41	0.49	0.5205	0.030	1.32
	81	0.508	0.5205	0.012	-

Table 5.6: Example 5.6: Convergence Rates for Flux: Cell-Node and Cell-Surface Discretization





a) Cell-Surface Discretization

b) Cell-Node Discretization

Figure 5.25: Example 5.6: Pressure Isolines For Sand-Shale Problem

computing the stream function and then drawing its isolines, are shown in Figure 5.24. The isolines of the pressure are presented in Figure 5.25, which shows the more regular behavior of the pressure for the case of the cell-surface discretization. The convergence rates for the total flux are presented in Table 5.6. For N=M=21, the cell-node discretization gives absolutely unphysical result. This can be explained as follows: the fluxes are computed at nodes, and the corresponding elements of the matrix  ${\bf K}$  are also computed at the nodes by a harmonic average from four neighboring cells, so for  $20\times 20$  cells, all nodes have very small permeability, which almost blocks the system.

In is interesting to compare the accuracy of the our cell-surface discretization and mixed finite-element method, described in Durlofsky [6], as function of number of unknowns. In [6], the total flux, which is obtained using 1240 unknowns, is 0.4508, while our method gives a flux of 0.451 for M=N=21 (the correct result is 0.5205). For a uniform grid used in this example, and for a diagonal  $\mathbf K$  matrix, as used in this example, the fluxes can be eliminated from the system and then the resulting system of linear equations contains only the pressure. That is, the number of unknowns, taking into account the Dirichlet boundary conditions on the left and right boundaries, is equal to  $(M-1)\times (N+1)=440$ . If the fluxes are not eliminated, then the number of unknowns is equal to  $(M-1)\times (N+1)+M\times (N-1)+(N-2)\times (M-1)=1240$ , exactly as for the mixed finite-element method. Because of these differences, to compare the two methods, it is not appropriate to compare only the number of unknowns, but the structure of the matrix and the solution procedure for the system of linear equations must also be considered.

#### 5.7 Flow Through a System Containing an Impermeable Streak

This example, similar to one in *Durlofsky* [6], uses the logically rectangular grid on the unit square shown Figure 5.26. The top curve is chosen to be an arc of a circle with the

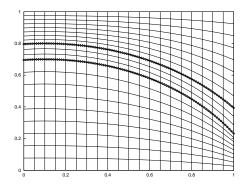


Figure 5.26: Example 5.7: The Geometry and the Grid for the Streak

center at (0.1, -0.4) and radius equal to 1.2. The bottom curve an arc of a circle with the same center and with radius equal to 1.1.

The permeability throughout the domain is uniform and isotropic ( $\mathbf{K} = \mathbf{I}$ ), except in the low-permeability streak where the permeability is set such that the component parallel to the local streak orientation ( $\mathbf{k}_{\parallel}$ ) is equal to 0.1 and the component perpendicular to the streak orientation ( $\mathbf{k}_{\perp}$ ) is equal to 0.001. In the streak, the tensor  $\mathbf{K}$  is a full tensor, in terms of its Cartesian components, which vary with (x,y) and are readily determined from the knowledge of  $\mathbf{k}_{\parallel}$  and  $\mathbf{k}_{\perp}$ . For the Cartesian components  $\mathbf{K}_{xx}, \mathbf{K}_{xy}, \mathbf{K}_{yy}$ , which are used in cell-node discretization, the transformation formulas are:

$$K_{xx} = \mathbf{k}_{\parallel} \cos^2 \varphi + \mathbf{k}_{\perp} \sin^2 \varphi \,, \tag{5.21}$$

$$K_{xy} = (\mathbf{k}_{\parallel} - \mathbf{k}_{\perp}) \cos \varphi \sin \varphi , \qquad (5.22)$$

$$K_{yy} = \mathbf{k}_{||} \sin^2 \varphi + \mathbf{k}_{\perp} \cos^2 \varphi \,, \tag{5.23}$$

where  $\varphi = \varphi(x,y)$  is the angle of rotation of the orthogonal coordinate system where the tensor **K** is diagonal and has components  $k_{\parallel}$  and  $k_{\perp}$ . In our case

$$\sin \varphi = -\frac{x'}{\sqrt{(x')^2 + (y')^2)}},$$
 (5.24)

$$\cos \varphi = \frac{y'}{\sqrt{(x')^2 + (y')^2}},$$
 (5.25)

where x' = x - 0.1 and y' = y + 0.4.

Figure 5.27 displays the velocity field for case of the cell-node discretization (the length of arrows are proportional to module of the vectors). The results of the cell-surface discretization will be similar. As expected physically, no flow enters the streak, so these results are qualitatively similar to the best results in *Durlofsky* [6].

#### 5.8 Non-Diagonal, Piecewise Continuous K

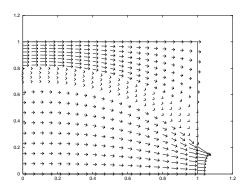


Figure 5.27: Example 5.7: Velocity Field for Cell-Node Discretization

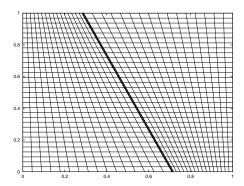


Figure 5.28: Example 5.8: Grid and Discontinuity Line, M=N=33

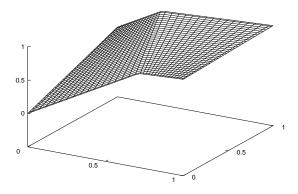


Figure 5.29: Example 5.8: Pressure, M = N = 33

In this example from Das, Schaffer, Steinberg, and Weber [18], the region is the unit square and the boundary conditions are the normal flux given by the exact solution on the top and the bottom boundaries, and Dirichlet condition given by the exact solution on the left and right boundaries. The non-diagonal permeability (or diffusion) matrix has a jump discontinuity of height  $\lambda$  along the line  $rx + sy = \delta$ , where  $0 \le r, s, \delta \le 1$  and r + s = 1 (see Figure 5.28). The matrix **K** is

$$\mathbf{K} = \mathbf{k}(x, y) \begin{pmatrix} 1 & 1/10 \\ 1/10 & 1 \end{pmatrix},$$
 (5.26)

where

$$k(x,y) = \begin{cases} 1, & \text{if } 0 \le r \, x + s \, y < \delta, \\ \lambda, & \text{if } \delta \le r \, x + s \, y < 1, \end{cases}$$
 (5.27)

The exact solution for the case when the right-hand side equals zero is

$$u(x,y) = \begin{cases} \frac{\lambda (r x + s y)}{1 + \delta (\lambda - 1)}, & \text{if } 0 \leq r x + s y < \delta, \\ \frac{(r x + s y) + \delta (\lambda - 1)}{1 + \delta (\lambda - 1)}, & \text{if } \delta \leq r x + s y < 1, \end{cases}$$

$$(5.28)$$

As expected, the method for the case of cell-surface discretization is exact for the piecewise linear solution.

Figure 5.29 displays the approximate solution for r = 0.7 and  $\lambda = 10$ , for the cell-node discretization. Table 5.7 verifies that the convergence rates are second-order.

M	max	$L_2$ -norm	$q_{max}$	$q_2$
	norm			
17	5.98E-4	2.15E-4	1.74	1.99
33	1.78E-4	5.38E-5	1.83	1.98
65	4.98E-5	1.36E-5	-	-

Table 5.7: Example 5.8: Convergence Rates: Cell-Node Discretization

## Acknowledgment

This work performed under the auspices of the US Department of Energy under contract W-7405-ENG-36 and DOE/BES Program in the Applied Mathematical Sciences Contract KC-07-01-01.

The authors thank J.E. Dendy, L.G. Margolin, and J.E. Morel for many fruitful discussions and comments on various drafts of the paper, and B. Swartz for discussions on the form of the conductivity tensor.

# A The Truncation Error for the Discrete Analog of K grad

In this appendix, the approximation properties of **K** grad are defined in terms of projection operators for scalar- and vector-valued continuum functions. For the cell-surface discretization, scalar valued functions are projected onto their values at cell centers:

$$(p_h(u))_{(i,j)} = u(x_{(i,j)}^c, y_{(i,j)}^c),$$
 (A.1)

where

$$x_{(i,j)}^{c} = \frac{x_{(i,j)} + x_{(i+1,j)} + x_{(i+1,j+1)} + x_{(i,j+1)}}{4},$$

$$y_{(i,j)}^{c} = \frac{y_{(i,j)} + y_{(i+1,j)} + y_{(i+1,j+1)} + y_{(i,j+1)}}{4},$$
(A.2)

and vector-valued functions are projected onto the normal component of the vector in the middle of the sides of the cells:

$$P_h = \left(egin{array}{c} P \, \xi_h \ P \, \eta_h \end{array}
ight) \, , \qquad \qquad (\mathrm{A.3})$$

where

$$(P\xi_h(\vec{w}))_{(i,j)} = (\vec{w}, \vec{n})|_{(x\xi,y\xi)_{(i,j)}}, \quad (P\eta_h(\vec{w}))_{(i,j)} = (\vec{w}, \vec{n})|_{(x\eta,y\eta)_{(i,j)}}, \quad (A.4)$$

and

$$x\xi_{(i,j)} = \frac{x_{(i,j)} + x_{(i,j+1)}}{2}, \quad y\xi_{(i,j)} = \frac{y_{(i,j)} + y_{(i,j+1)}}{2},$$

$$x\eta_{(i,j)} = \frac{x_{(i,j)} + x_{(i+1,j)}}{2}, \quad y\eta_{(i,j)} = \frac{y_{(i,j)} + y_{(i+1,j)}}{2}.$$
(A.5)

Then the truncation error for discrete operator  $\mathcal{G}$  is defined to be

$$\psi_{\mathcal{G}}(u) = \mathcal{G}(p_h u) - P_h(\mathbf{K} \operatorname{\mathbf{grad}} u).$$
 (A.6)

Our goal is to show that for piecewise linear function u and general grids, the discrete operator  $\mathcal{G}$  approximating  $\mathbf{K}$  grad, is exact on linear functions. We restrict the proof to the case of a diagonal matrix  $\mathbf{K} = \mathbf{k}(x,y)\mathbf{I}$  with piecewise constant function  $\mathbf{k}(x,y)$ . The results also hold for non-diagonal and piecewise constant  $\mathbf{K}$  and the proof of the general result is similar to the one given. Also, we assume that the discontinuities of the coefficient  $\mathbf{K}$  coincide with curve where u may not be differentiable, and that this curve is continuous and aligned with the cell boundaries.

Because the function u is piecewise linear, the derivatives  $\partial u/\partial x$ , and  $\partial u/\partial y$  are piecewise constants. Because the flux is continuous at interfaces, the Cartesian components of flux,

$$GX = k \frac{\partial u}{\partial x}, \quad GY = k \frac{\partial u}{\partial y},$$
 (A.7)

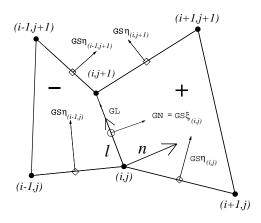


Figure A.30: The Local Basis System

are constant over the entire domain.

The definition of the components of the vector  $\vec{G} = (GS\xi, GS\eta) = \mathbf{k} \operatorname{\mathbf{grad}} u$  are

$$GS\xi_{(i,j)} = GX \, ns \xi_{(i,j)}^x + GY \, ns \xi_{(i,j)}^y \,,$$
 (A.8)  
 $GS\eta_{(i,j)} = GX \, ns \eta_{(i,j)}^x + GY \, ns \eta_{(i,j)}^y \,,$ 

where  $(ns\xi^x, ns\xi^y)$  are components of the normal to side  $S\xi$ , and  $(ns\eta^x, ns\eta^y)$  are components of the normal to side  $S\eta$ .

For this investigation, it is convenient to introduce the local basis system with its origin at the node  $(x_{(i,j)}, y_{(i,j)})$ , axis l directed along the side  $S\xi_{(i,j)}$ , and axis n is perpendicular to this side (see, Figure A.30). The projections of a vector  $\vec{G}$  on the axes of this local orthogonal coordinate system are denoted  $GN = GS\xi_{(i,j)}$  and GL. Therefore, GN is the component normal to the interface, and GL is the component tangential to the interface. These components can be written in terms of GX and GY and vice versa:

$$GN = GX \, ns \xi_{(i,j)}^{x} + GY \, ns \xi_{(i,j)}^{y},$$

$$GL = -GX \, ns \xi_{(i,j)}^{y} + GY \, ns \xi_{(i,j)}^{x},$$

$$GX = GN \, ns \xi_{(i,j)}^{x} - GL \, ns \xi_{(i,j)}^{y},$$

$$GY = GN \, ns \xi_{(i,j)}^{y} + GL \, ns \xi_{(i,j)}^{x}.$$
(A.9)

Using (A.8), and (A.9),  $GS\xi$  and the components  $GS\eta$  on all sides of the cells (i,j) and (i-1,j) can be express in terms of GN and GL as follows (see Figure A.30):

$$\begin{split} GS\xi_{(i,j)} &= GN \;, \\ GS\eta_{(i,j)} &= \sin\varphi_{(i,j)}^{(i,j)} GL - \cos\varphi_{(i,j)}^{(i,j)} GN \;, \\ GS\eta_{(i,j+1)} &= \sin\varphi_{(i,j+1)}^{(i,j)} GL + \cos\varphi_{(i,j+1)}^{(i,j)} GN \;, \end{split} \tag{A.10}$$

$$GS\eta_{(i-1,j+1)} = \sin arphi_{(i,j+1)}^{(i-1,j)} GL - \cos arphi_{(i,j+1)}^{(i-1,j)} GN \ , \ GS\eta_{(i-1,j)} = \sin arphi_{(i,j)}^{(i-1,j)} GL + \cos arphi_{(i,j)}^{(i-1,j)} GN \ .$$

The definition (A.6) of the truncation error  $\psi_{\mathcal{G}}$  and the definition (4.48) of the operator  $\mathcal{G}$  gives

$$\mathcal{S} \, \psi_{\mathcal{G}} = \mathcal{S} \, \left( P_h(\mathbf{K} \, \mathbf{grad} \, u) \right) - \mathcal{D}^\dagger \, \mathcal{M} \, \left( p_h \, u \right) \, .$$
 (A.11)

The right-hand side of this equation is estimated by estimating each term.

To estimate the first term in (A.11), let  $\tilde{G} = \mathcal{S} (P_h(\mathbf{k} \operatorname{\mathbf{grad}} u))$  and then

$$\begin{split} \tilde{GS}\xi_{(i,j)} &= \\ \frac{1}{k_{(i,j)}^{+}} \left[ \left( \frac{V_{(i,j)}^{(i,j)}}{\sin^{2}\varphi_{(i,j)}^{(i,j)}} + \frac{V_{(i,j+1)}^{(i,j)}}{\sin^{2}\varphi_{(i,j+1)}^{(i,j)}} \right) GS\xi_{(i,j)} + \right. \\ \frac{V_{(i,j)}^{(i,j)} \cos\varphi_{(i,j)}^{(i,j)}}{\sin^{2}\varphi_{(i,j)}^{(i,j)}} GS\eta_{(i,j)} - \frac{V_{(i,j+1)}^{(i,j)} \cos\varphi_{(i,j+1)}^{(i,j)}}{\sin^{2}\varphi_{(i,j+1)}^{(i,j)}} GS\eta_{(i,j+1)} \right] + \\ \frac{1}{k_{(i,j)}^{-}} \left[ \left( \frac{V_{(i,j)}^{(i-1,j)}}{\sin^{2}\varphi_{(i,j)}^{(i-1,j)}} + \frac{V_{(i,j+1)}^{(i-1,j)}}{\sin^{2}\varphi_{(i,j+1)}^{(i-1,j)}} \right) GS\xi_{(i,j)} - \right. \\ \frac{V_{(i,j)}^{(i-1,j)} \cos\varphi_{(i,j)}^{(i-1,j)}}{\sin^{2}\varphi_{(i,j)}^{(i-1,j)}} GS\eta_{(i-1,j)} + \frac{V_{(i-1,j+1)}^{(i,j)} \cos\varphi_{(i,j+1)}^{(i-1,j)}}{\sin^{2}\varphi_{(i,j+1)}^{(i-1,j)}} GS\eta_{(i-1,j)} \right] . \end{split}$$

Because

$$V_{(i,j+l)}^{(i-k,j)} = 0.25 \, S\xi_{(i,j)} \, S\eta_{(i-k,j+l)} \, \sin\varphi_{(i,j+l)}^{(i-k,j)} \,, \tag{A.13}$$

Equation (A.12) can be transformed to

$$\begin{split} \tilde{GS}\xi_{(i,j)} &= \\ 0.25\,S\xi_{(i,j)} \left\{ \frac{1}{\mathbf{k}^{+}} \left[ \left( \frac{S\eta_{(i,j)}}{\sin\varphi_{(i,j)}^{(i,j)}} + \frac{S\eta_{(i,j+1)}}{\sin\varphi_{(i,j+1)}^{(i,j)}} \right) GS\xi_{(i,j)} + \right. \\ &\left. \frac{S\eta_{(i,j)} \cos\varphi_{(i,j)}^{(i,j)}}{\sin\varphi_{(i,j)}^{(i,j)}} GS\eta_{(i,j)} - \frac{S\eta_{i,j-1} \cos\varphi_{(i,j+1)}^{(i,j)}}{\sin\varphi_{(i,j+1)}^{(i,j)}} GS\eta_{(i,j+1)} \right] + \\ &\left. \frac{1}{\mathbf{k}^{-}} \left[ \left( \frac{S\eta_{(i-1,j)}}{\sin\varphi_{(i,j)}^{(i-1,j)}} + \frac{S\eta_{(i-1,j+1)}}{\sin\varphi_{(i,j+1)}^{(i-1,j)}} \right) GS\xi_{(i,j)} - \right. \\ &\left. \frac{S\eta_{(i-1,j)} \cos\varphi_{(i,j)}^{(i-1,j)}}{\sin\varphi_{(i,j)}^{(i-1,j)}} GS\eta_{(i-1,j)} + \frac{S\eta_{(i-1,j+1)} \cos\varphi_{(i,j+1)}^{(i-1,j)}}{\sin\varphi_{(i,j+1)}^{(i-1,j)}} GS\eta_{(i-1,j+1)} \right] \right\} \end{split}$$

Using the formulas in (A.10), the previous formula can be written in terms of GL and GN:

$$\tilde{GS}\xi_{(i,j)} = \tag{A.15}$$

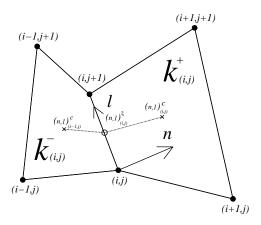


Figure A.31: One-sided Taylor Series

$$\begin{split} 0.25\,S\,\xi_{(i,j)}\,\left\{ \frac{1}{\mathbf{k}^{+}}\,\left[GN\,\left(S\eta_{(i,j)}\,\sin\varphi_{(i,j)}^{(i,j)}+S\eta_{(i,j+1)}\,\sin\varphi_{(i,j+1)}^{(i,j)}\right) + \right. \\ \left. GL\,\left(S\eta_{(i,j)}\,\cos\varphi_{(i,j)}^{(i,j)}-S\eta_{(i,j+1)}\,\cos\varphi_{(i,j+1)}^{(i,j)}\right)\right] + \\ \left. \frac{1}{\mathbf{k}^{-}}\,\left[GN\,\left(S\eta_{(i-1,j)}\,\sin\varphi_{(i,j)}^{(i-1,j)}+S\eta_{(i-1,j+1)}\,\sin\varphi_{(i,j+1)}^{(i-1,j)}\right) + \right. \\ \left. GL\,\left(-S\eta_{(i-1,j)}\,\cos\varphi_{(i,j)}^{(i-1,j)}+S\eta_{(i-1,j+1)}\,\cos\varphi_{(i,j+1)}^{(i-1,j)}\right)\right]\right\}\,. \end{split}$$

For the second term in (A.11), consider the  $\xi$  component of  $\mathcal{D}^{\dagger}$   $\mathcal{M}\left(p_{h}\left(u\right)\right)$  (see (4.52)), which is

$$S\xi_{(i,j)}\left[\left(p_h(u)\right)_{(i,j)}-\left(p_h(u)\right)_{(i-1,j)}\right]$$
 (A.16)

Because the function u is piece-wise linear, the truncation error is computed using one-sided Taylor series at the point  $(x\xi, y\xi)_{(i,j)}$ , which is the middle of side (i,j)-(i,j+1) (see Figure A.31). Values on the "right" of this side are labeled with a sub- or superscript "+", while values on the "left" of this side are labeled with sub- or superscript "-".

Because the local basis system (n, l) is used and the fact that the function u is linear on each side of the discontinuity,

$$egin{array}{lll} (p_h\,u)_{(i,j)} &=& u^+ + \left.rac{\partial u}{\partial n}
ight|_+ \, (n^c_{(i,j)} - n^{\xi}_{(i,j)}) + \left.rac{\partial u}{\partial l}
ight|_+ \, (l^c_{(i,j)} - l^{\xi}_{(i,j)})\,, \ & \ (p_h\,u)_{(i-1,j)} &=& u^- + \left.rac{\partial u}{\partial n}
ight|_- \, (n^c_{(i-1,j)} - n^{\xi}_{(i,j)}) + \left.rac{\partial u}{\partial l}
ight|_- \, (l^c_{(i-1,j)} - l^{\xi}_{(i,j)})\,. \end{array}$$

To simplify the notation, subscripts or superscripts for values at the center point  $(x\xi, y\xi)_{(i,j)}$  of the edge l are not used. Because GN and GL are constant over the entire domain, and the function u is continuous, that is,  $u^+ = u^- = u$  at the center point of the edge of l, the

previous equations become

$$(p_{h}(u))_{(i,j)} = u + \frac{GN}{k^{+}} (n_{(i,j)}^{c} - n_{(i,j)}^{\xi}) + \frac{GL}{k^{+}} (l_{(i,j)}^{c} - l_{(i,j)}^{\xi})$$

$$= u + \frac{1}{k^{+}} \left[ GN (n_{(i,j)}^{c} - n_{(i,j)}^{\xi}) + +GL (l_{(i,j)}^{c} - l_{(i,j)}^{\xi}) \right],$$

$$(p_{h}(u))_{(i-1,j)} = u + \frac{GN}{k^{-}} (n_{(i-1,j)}^{c} - n_{(i,j)}^{\xi}) + \frac{GL}{k^{-}} (l_{(i-1,j)}^{c} - l_{(i,j)}^{\xi})$$

$$= u + \frac{1}{k^{-}} \left[ GN (n_{(i-1,j)}^{c} - n_{(i,j)}^{\xi}) + +GL (l_{(i-1,j)}^{c} - l_{(i,j)}^{\xi}) \right].$$
(A.17)

And finally,

$$(p_{h}(u))_{(i,j)} - (p_{h}(u))_{(i-1,j)} =$$

$$\frac{1}{k^{+}} \left[ GN \left( n_{(i,j)}^{c} - n_{(i,j)}^{\xi} \right) + +GL \left( l_{(i,j)}^{c} - l_{(i,j)}^{\xi} \right) \right] +$$

$$\frac{1}{k^{-}} \left[ GN \left( n_{(i-1,j)}^{c} - n_{(i,j)}^{\xi} \right) + +GL \left( l_{(i-1,j)}^{c} - l_{(i,j)}^{\xi} \right) \right]$$

$$(A.18)$$

To compare (A.15) with (A.18), expressions for the coordinates of the vertices of the cells are needed in the local coordinate system:

$$\begin{split} &n_{(i,j)} = 0\,, \quad l_{(i,j)} = 0\,, \\ &n_{(i,j+1)} = 0\,, \quad l_{(i,j+1)} = S\xi_{(i,j)}\,, \\ &n_{(i+1,j)} = S\eta_{(i,j)}\,\sin\varphi_{(i,j)}^{(i,j)}\,, \quad l_{(i+1,j)} = S\eta_{(i,j)}\,\cos\varphi_{(i,j)}^{(i,j)}\,, \\ &n_{(i+1,j+1)} = S\eta_{(i,j+1)}\,\sin\varphi_{(i,j+1)}^{(i,j)}\,, \quad l_{(i+1,j+1)} = S\xi_{(i,j)} - S\eta_{(i,j+1)}\,\cos\varphi_{(i,j)}^{(i,j)}\,, \\ &n_{(i-1,j+1)} = -S\eta_{(i-1,j+1)}\,\sin\varphi_{(i,j+1)}^{(i-1,j)}\,, \quad l_{(i-1,j+1)} = S\xi_{(i,j)} - S\eta_{(i-1,j+1)}\,\cos\varphi_{(i,j+1)}^{(i-1,j)}\,, \\ &n_{(i-1,j)} = -S\eta_{(i-1,j)}\,\sin\varphi_{(i,j)}^{(i-1,j)}\,, \quad l_{(i-1,j)} = S\eta_{(i-1,j)}\,\cos\varphi_{(i,j)}^{(i-1,j)}\,, \end{split}$$

and also

$$n_{(i,j)}^{\xi} = 0$$
,  $l_{(i,j)}^{\xi} = 0.5 S \xi_{(i,j)}$ . (A.19)

These expression give

$$\begin{split} n_{(i,j)}^c - n_{(i,j)}^\xi &= \frac{1}{4} \left( 0 + 0 + S \eta_{(i,j)} \sin \varphi_{(i,j)}^{(i,j)} + S \eta_{(i,j+1)} \sin \varphi_{(i,j+1)}^{(i,j)} \right) - 0 \\ &= \frac{1}{4} \left( S \eta_{(i,j)} \sin \varphi_{(i,j)}^{(i,j)} + S \eta_{(i,j+1)} \sin \varphi_{(i,j+1)}^{(i,j)} \right) \;, \\ l_{(i,j)}^c - l_{(i,j)}^\xi &= \frac{1}{4} \left( 0 + S \xi_{(i,j)} + S \eta_{(i,j)} \cos \varphi_{(i,j)}^{(i,j)} + S \xi_{(i,j)} + S \eta_{(i,j+1)} \cos \varphi_{(i,j+1)}^{(i,j)} \right) - \frac{1}{2} S \xi_{(i,j)} \\ &= \frac{1}{4} \left( S \eta_{(i,j)} \cos \varphi_{(i,j)}^{(i,j)} - S \eta_{(i,j+1)} \cos \varphi_{(i,j+1)}^{(i,j)} \right) \;. \end{split}$$

From these formulas it is clear that the coefficients of  $1/k^+$  in the expressions  $\mathcal{S}$   $P_h(\mathbf{k}$  grad u) and  $\mathcal{D}^{\dagger}\mathcal{M}$   $(p_h\,u)$  are the same. Similar considerations are valid for  $n_{(i-1,j)}^c-n_{(i,j)}^\xi$  and  $l_{(i-1,j)}^c-l_{(i,j)}^\xi$ . Therefore

$$S\psi_{\mathcal{G}} = 0, \tag{A.20}$$

and S is invertible, so consequently  $\psi_{\mathcal{G}}=0$ , that is, the operator  $\mathcal{G}$  is exact for piecewise linear functions.

#### References

- [1] M. H. Carpetner, D. Gottlieb, S. Abarbanel, Time Stable Boundary Conditions for Finite-Difference Schemes Solving Hyperbolic Systems: Methodology and Application to High-Order Compact Schemes, Journal of Computational Physics, 111, pp. 220-236 (1994).
- [2] P. I. Crumpton, G. I. Shaw, and A. F. Ware. Discretization and Multigrid Solution of Elliptic Equations with Mixed Derivative Terms and Strongly Discontinuous Coefficients, J. Comput. Phys., 116, (1995), pp. 343-358.
- [3] R. E. Alcouffe, A. Brandt, J. E. Dendy, Jr., and J. W. Painter, *The Multi-grid Method for the Diffusion Equation with Strongly Discontinuous Coefficients*, SIAM J. Sci. Stat. Comp., 2, (1981), pp. 430-454.
- [4] J. E. Dendy, Jr., Black Box Multigrid, J. Comp. Phys., 48, (1982), pp. 366-386.
- [5] L. J. Durlofsky, A Triangle Based Mixed Finite Element-Finite Volume Technique for Modeling Two Phase Flow through Porous Media, Journal of Computational Physics, 105, pp. 252-266, (1993).
- [6] L. J. Durlofsky, Accuracy of Mixed and Control Volume Finite Element Approximations to Darcy Velocity and Related Quantities, Water Resources Research, Vol. 30, No. 4, pp. 965-973, April 1994.
- [7] V. V. Jikov, S. M. Kozlov and O. A. Oleinik, Homogenization of Differential Operators and Integral Functionals, Springer-Verlag, 1994, p.37.
- [8] D.S. Kershaw, The Incomplete Choleski-Conjugate Gradient Methods for the Iterative Solution of Systems of Linear Equations, J. Comp. Phys., 26, pp. 43-65, (1978).
- [9] P.M. Knupp and S. Steinberg, *The Fundamentals of Grid Generation*, CRC Press, Boca Raton, 1993.
- [10] S. K. Lele, Compact Finite Difference Schemes with Spectral-like Resolution, Journal of Computational Physics, 103, pp. 16-42, (1992).
- [11] S. H. Leventhal, An Operator Compact Implicit Method of Exponential Type, Journal of Computational Physics, 46, pp. 138-165, (1982).
- [12] R. J. MacKinnon and G. F. Carey. Analysis of Material Interface Discontinuities and Superconvergent Fluxes in Finite Difference Theory, J. Comput. Phys., 75, pp. 151-167, (1988).
- [13] L. G. Margolin and J. J. Pyun, A Method for Treating Hourglass Pattern, LA-UR-87-439, Report of Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

- [14] A. Favorskii, A. Samarskii, M. Shashkov and V. Tishkin Operational Finite-Difference Schemes, Differential Equations, 17, (1981), 854-862.
- [15] M. Shashkov, and S. Steinberg, A New Numerical Method for Solving Diffusion Equations with Rough Coefficients in Rough Grids, Submitted to Journal of Computational Physics.
- [16] M. Shashkov, and S. Steinberg, Support-Operator Finite-Difference Algorithms for General Elliptic Problems, Journal of Computational Physics, 118, pp. 131-151, (1995).
- [17] M. Shashkov, Conservative Finite-Difference Methods on General Grids, CRC Press, Boca Raton, 1995.
- [18] B. Das, S. Schaffer, S. Steinberg, and S. Weber Finite Difference Methods for Modeling Porous Media Flows, Transport in Porous Media, 17, (1994), pp. 171-200.
- [19] de Zeeuw, P. M. Matrix-dependent Prolongation and Restriction in the Blackbox Multigrid Solver, J. Comput. Appl. Math., 3, (1990), pp. 1-27.
- [20] P. van Beek, R. R. P. van Nooyen and P. Wesseling, Accurate Discretization of Gradients on Non-uniform Curvilinear Staggered Grids, Journal of Computational Physics, 117, pp. 364-367, (1995).
- [21] R. S. Bernard and H. Kapitza, How to Discretize the Pressure Gradient for Curvilinear MAC Grids, Journal of Computational Physics, 99, pp. 288-298, (1992).
- [22] M. F. Wheeler and R. Gonzalez, Mixed Finite Element Methods for Petroleum Reservoir Engineering Problems, Computing Methods in Applied Science and Engineering, VI, Eds. R. Glowinski and J.-L. Lions. pp. 639-657.
- [23] J.M. Hyman, R.J. Knapp, and J.C. Scovel, High Order Finite Volume Approximations of Differential Operators on Nonuniform Grids, Physica D 60 (1992), 112-138.
- [24] A. Favorskii, T. Korshiya, M. Shashkov and V.Tishkin, Variational Approach to The Construction of Finite-Difference Schemes for The Diffusion Equations for Magnetic Field. Differential equations, 1982, Vol.18, N 7, pp.863-872.
- [25] D. S. Kershaw. Differencing of the Diffusion Equation in the Lagrangian Hydrodynamic Codes, Journal of Computational Physics, 39, pp. 375-395, (1981).
- [26] G. Knorr, G. Joyce and A.J. Marcus, Fourth Order Poisson Solver for Simulation of Bounded Plasmas, J. Comp. Phy., 38 (1980), 227-236.
- [27] J. M. Morel, J. E. Dendy, Jr., Michael L. Hall, and Stephen W. White. A Cell-Centered Lagrangian-Mesh Diffusion Differencing Scheme, Journal of Computational Physics, 103, pp. 286-299, (1992).

- [28] T. Robey, em Development of Models for Fast Fluid Pathways Through Unsaturated Heterogeneous Porous Media. Yucca Mountain Site Characterization Project, Contractor Report, SAND93-7109, Sandia National Laboratories, Albuquerque, New Mexico.
- [29] T. Robey, An Adaptive Grid Technique for Minimizing Heterogeneity of Cells or Elements, Preprint, Spectra Research Institute, Albuquerque, NM, 1994.
- [30] I. Babuska, G. Calos and E. Osborn. Special Finite Element Methods for Class of Second Order Elliptic Problems with Rough Coefficients, SIAM J. Numer. Anal., Vol. 31, No. 4, (1994), pp. 945-981.
- [31] R. Leveque and Z. Li. The Immersed Interface Method for Elliptic Equations with Discontinuous Coefficients and Singular Sources, SIAM J. Numer. Anal., Vol. 31, No. 4, (1994), pp. 1019-1044.
- [32] T. Arbogast, P. T. Keenan, M. F. Wheeler, and I. Yotov. Logically Rectangular Mixed Methods for Darcy Flow on General Geometry, in Proceedings of "Symposium on Reservoir Simulation", 12-15 February 1995, San Antonio, TX, Society of Petroleum Engineers, Inc., pp. 51-59.
- [33] M. G. Edwards. Symmetric Flux Continuous Positive Definite Approximation of Elliptic Full Tensor Pressure Equation in Local Conservation Form, in Proceedings of "Symposium on Reservoir Simulation", 12-15 February 1995, San Antonio, TX, Society of Petroleum Engineers, Inc., pp. 553-562.